Transcriptional Profiling of Plant Embryogenesis.

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Durham

by

Matthew William Beresford Spencer BSc (dunelm)
School of Biological and Biomedical Sciences
University of Durham

September 2005

Abstract

The process of embryogenesis in higher plants is a critical stage of the sporophytic life cycle, transforming the fertilised egg cell via a precise sequence of events into a multi-cellular organism. It is during embryogenesis that the body plan of the developing plant is established. Analysis of transcriptional changes occurring during the establishment of the developing plants body plan has been limited due to the technical difficulties associated with accessing the developing embryo.

This thesis demonstrates the application of laser capture microdissection to the analysis of embryogenesis in the model plant *Arabidopsis thaliana*. This technique has been used in combination with DNA microarray technology to allow a global analysis of gene expression in the cotyledon, root and shoot apical meristem regions of the torpedo-stage embryo.

Validation of the approach has been achieved by comparison of the ATH1 GeneChip® data obtained, with published gene expression patterns confirmed by *in situ* hybridisation and promoter::GUS analysis. Further validation was successfully undertaken through the creation of promoter::GUS constructs for a number of previously uncharacterised putative transcription factor genes, selected on the basis of differential expression between the cotyledon and root regions. Initial attempts to assign putative function to these genes through an analysis of T-DNA insertion lines yielded no aberrant phenotypes.

Transcriptional profiling of embryogenesis from the globular-stage through to the torpedo-stage was carried out using GeneSpring, uncovering distinct spatial and temporal expression patterns, and revealing a number of genes of potential interest for further research.

Declaration

The material contained in this thesis has not been submitted for a degree in this or any other University.

Statement of Copyright

In

Loving Memory

Of

Granny

Acknowledgements

First and foremost I wish to thank my supervisor, Keith Lindsey for all his advice and encouragement, for always having confidence in me, and for critically reading this manuscript. I wish to acknowledge the BBSRC and Syngenta for funding the project.

Many thanks to lab members past and present. Gill, for showing me that doing a PhD and always being happy aren't mutually exclusive!; Stu, for keeping my ice bucket topped up (despite occasionally missing); Paul, for having such fine taste in music; Jen, Mags, Marta and Roya for all the great advice and support. Elerie, Nick and Lucy for making me wish I was staying in the lab a little bit longer. I would also like to thank all the friendly and helpful people in the department who have gone out of their way to help me over the years.

Thanks also to all the people down at NASC for looking after me so well, particularly to Sean May for allowing me to use their GeneSpring license and to Martin Broadley for all his help getting started.

To all my long-suffering housemates and good friends, Simon, Vicki and James; Shaz, Heather and Woody and last but by no means least (otherwise I would be in trouble!) Stu.

I would like to thank my two best advisors and friends, my favourite married couple; Samuel Duncan (the devil!!!) and Rebecca Nesbit (my angel). Everyone in Castle, particularly all the fantastic people I have had the pleasure of knowing in the basketball and boat clubs.

To my family, Mum, Dad, Richard and Robert, without whose constant love and support I would certainly not be who I am today or achieved so much, I give this work to you.

Lastly, I would like to thank Beaker for being such a good role model and Super Matty for finally coming to the rescue.

Publications

Publications resulting from work carried out for this thesis:

Research Papers

Stuart Casson, Matthew Spencer, Katherine Walker, Keith Lindsey (2005) Lasercapture microdissection for the analysis of gene expression during embryogenesis of *Arabidopsis*. *The Plant Journal* **42**, 111-123.

Matthew W.B. Spencer, Stuart A. Casson, Keith Lindsey. Transcriptional profiling of plant embryogenesis (in preparation).

Conference Abstracts

Stuart A. Casson, Matthew W. Spencer, Keith Lindsey (2003). Laser-capture microdissection to analyse gene transcription in *Arabidopsis* embryos. *Comp. Biochem. Physiol.* **134/A Suppl.** S199.

Matthew W. Spencer, Stuart A. Casson, Keith Lindsey (2004). Exploiting laser capture microdissection to elucidate spatial gene expression during plant embryogenesis. *Abstr. SEB poster, Heriot-Watt meeting.*

Matthew Spencer, Stuart Casson, Keith Lindsey (2004). Exploiting laser capture microdissection to elucidate spatial gene expression during plant embryogenesis. *Abstr. GARNet Meeting, Leicester.*

Keith Lindsey, Stuart A. Casson, Matthew W. Spencer, Katherine Walker (2004). The virtual embryo: modelling *Arabidopsis* embryogenesis. *Abstr. Mathematical Modelling of Plant Development and Gene Networks*, University of Warwick.

Matthew Spencer, Stuart Casson, Keith Lindsey (2004). Exploiting laser capture microdissection to elucidate spatial gene expression during plant embryogenesis. T10-038. *Abstr. 15th Int. Conf. Arabidopsis Res., Berlin.*

Keith Lindsey, Stuart Casson, Matthew Spencer. (2005). Laser capture microdissection for high resolution transcriptional profiling. *Abstr. GARNet Meeting, Norwich.* L07.

Abbreviations

ABA	Abscisic acid
ABI3	*ABSCISIC ACID-INSENSITIVE3*
Ac/Ds	Activator/dissociator element system
AGO	*ARGONAUTE*
ANT	*AINTEGUMENTA*
AP2	Apetala-2 transcription factor family
APC	anaphase-promoting complex
AREs	auxin response elements
ARF	*AUXIN RESPONSE FACTOR*
aRNA	amplified or antisense RNA
AS1	*ASYMMETRIC LEAVES 1*
AtLTP1	*LIPID TRANSFER PROTEIN 1*
AtML1	*Arabidopsis thaliana MERISTEM LAYER 1*
AXR	*AUXIN RESISTANT*
BFA	fungal toxin brefeldin A

BDL	*BODENLOS*
bp	base pair
BR	brassinosteroid hormones
BSA	bovine serum albumin
CAL	*CAULIFLOWER*
CaMV35S	Cauliflower mosaic virus 35S promoter
CAPS	Cleaved Amplified Polymorphic Sequences
cDNA	complementary DNA
CLV	*CLAVATA*
cm	centimetre
cRNA	complementary RNA
CTS	*COMATOSE*
CUC	*CUP-SHAPED COTYLEDON*
Cy3/5	Fluorescent cyanine dye
dATP	2'-deoxyadenosine 5'-triphosphate
DEPC	diethylpyrocarbonate

DNA	deoxyribonucleic acid
DNase	deoxyribonuclease (an enzyme that breaks down DNA)
dNTPs	2'-deoxynucleoside 5'-triphosphates
DPG	days post germination
DRL1	*DEFORMED ROOTS AND LEAVES1*
DTT	dithiothreitol
dTTP	2'-deoxythymidine 5'-triphosphate
dUTP	2'-deoxyuridine 5'-triphosphate
EDTA	ethylene diamine tetraacetic acid, disodium salt
EMB30	*EMBRYO DEFECTIVE 30*
EMS	ethyl methane sulphonate
ESTs	Expressed Sequence Tag
FIL	*FILAMENTOUS FLOWER*
FK	*FACKEL*
FUS3	*FUSCA3*
GA	gibberellic acid

gDNA	genomic DNA
GEF	guanine-nucleotide exchange factor
GFP	green fluorescent protein
GN	*GNOM*
GUS	β-glucuronidase
gusA (uidA)	gene encoding β-glucuronidase
HBT	*HOBBIT*
HCl	hydrochloric acid
HEPES	N-(2'-hydroxyethyl)-piperazine-N'-2-ethanesulfonic acid
HRGP's	hydroxyproline rich glycoproteins
IAA	indole-3-acetic acid
IPCR	Inverse polymerase chain reaction
IR	infrared
KAN	*KANADI*
Kb	kilobase (length of 1000 nucleotides)
KV	Kilovolt

KNOX	*KNOTTED1*-like homeobox
KOH	potassium hydroxide
lacZ	*β galactosidase*
LB	Luria-Bertani medium
LCM	Laser capture microdissection
LEA	Late Embryogenesis Abundant
LEC1	*LEAFY COTYLEON1*
LPC	Laser Pressure Catapulting
MAPKK	Mitogen-activated protein kinase kinase
MAS5	Microarray suite 5.0 software
Mb	Megabase (length of a million nucleotides)
µF	microfarad
Mg	magnesium
$MgCl_2$	magnesium chloride
MIPS	Munich information centre for protein sequences
ml	millilitre

µm	micrometre (micron)
mM	millimole
MM	'mismatch'
MOPS	4-Morpholinepropanesulfonic acid
MOR1	*MICROTUBULE ORGANIZATION1*
MP	*MONOPTEROS*
mRNA	messenger RNA
MS	Murashige and Skoog medium
MS2	*Male Sterile 2*
NaCl	sodium chloride
NAM	*NO APICAL MERISTEM*
NASC	Nottingham Arabidopsis Stock Centre
ng	nanogram (10^{-9} g)
NOS	nopaline synthase
NPTII	gene encoding neomycin phosphotransferase
nt	nucleotide

Ω	ohm
OD	optical density
PALM	position ablative laser microbeam
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pg	picogram
PHB	*PHABULOSA*
PHAN	*PHANTASTICA*
PHV	*PHAVOLUTA*
PKL	*PICKLE*
PID	*PINOID*
PIN	PIN FORMED
PM	'perfect-match'
PNH	*PINHEAD*
REV	*REVOLUTA*
RMA	Robust multi-array average

RNA	ribonucleic acid
RNAi	RNA Interference
RNase	ribonuclease (an enzyme that breaks down RNA)
rpm	revolutions per minute
RS2	*ROUGH SHEATH2*
RT-PCR	Reverse transcription PCR
SAM	shoot apical meristem
SCF	Ubiquitin ligase complex (SKP1, CUL1 and F-Box protein)
sdH_2O	sterile (autoclaved) distilled water
SDS	sodium dodecyl sulphate
SE	*SERRATE*
sec	Seconds
STM	*SHOOT MERISTEMLESS*
TAE	tris-acetate EDTA buffer
TAIR	The *Arabidopsis* information resource
T-DNA	transferred DNA

TIGR	The institute for genomic research
Ti plasmid	tumour-inducing plasmid
TTN	*TITAN*
Tris	tris(hydroxymethyl)-aminomethane
TUA4	*α-TUBULIN4*
twn	*twin* mutant
var.	variety
vir	*virulence* region
VPI	VIVIPAROUS-1
v/v	By volume
WOX	*WUS-RELATED HOMEOBOX*
WUS	*WUSCHEL*
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
X-Gluc	5-bromo-4-chloror-3-indoyl β-D-glucuronic acid
YAB	*YABBY*
YDA	*YODA*

ZLL ZWILLE

Contents

APPENDICES

List of Figures and Tables

Figures

Tables

4.5 *EMB* genes determined to show spatial differential expression during the torpedo-stage of embryogenesis.

5.1 The ten most differentially expressed significant genes up-regulated in the apical region of the globular stage embryo compared to the basal region.

5.2 The ten most differentially expressed significant genes up-regulated in the basal region of the globular stage embryo compared to the apical region.

5.3 The ten most differentially expressed significant genes up-regulated in the cotyledon region of the heart stage embryo compared to the root region.

5.4 The ten most differentially expressed significant genes up-regulated in the root region of the heart stage embryo compared to the cotyledon region.

5.5 The fifty most differentially expressed significant genes up-regulated in the cotyledon region of the torpedo stage embryo compared to the root region.

5.6 The fifty most differentially expressed significant genes up-regulated in the root region of the torpedo stage embryo compared to the cotyledon region.

5.7 The ten most differentially expressed significant genes up-regulated in the meristem of the torpedo stage embryo compared to the apical region of the globular stage embryo.

5.8 The ten most differentially expressed significant genes up-regulated in the apical region of the globular stage embryo compared to the meristem of the torpedo stage embryo.

5.9 The ten most differentially expressed significant genes up-regulated in the cotyledon of a 7dpg seedling compared to the root.

5.10 The ten most differentially expressed significant genes up-regulated in the root of a 7dpg seedling compared to the cotyledon.

5.11 A comparison of differentially expressed genes, between tissues and developmental stages.

Chapter 1

Introduction

1.0 Introduction

The process of embryogenesis in higher plants is a critical stage of the sporophytic life cycle, transforming the fertilised egg cell via a precise sequence of events into a multi-cellular organism. Embryogenesis in higher animals usually gives rise to what is effectively a miniature form of the adult with no deviation in post-embryonic development from the established body plan. In contrast to the higher animals, embryogenesis in higher plants generates a juvenile as opposed to miniature form of the adult (Berleth and Chatfield, 2002). The establishment of shoot and root stem-cell systems (meristems) provides the capacity for the increasingly complex architecture of post-embryonic development, which gives rise to the species-specific characteristics of the adult plant. The primary shoot meristem is the source of all above-ground organs generated post-embryonically, maintaining a fine balance between proliferation of the stem-cell population and differentiation. The primary root meristem enables the primary root to grow through extension (Fletcher, 2002; Bäurle and Laux, 2003; Jiang and Feldman, 2005).

The growth and development of higher plants can be seen to occur along two axes: the apical-basal axis and the radial axis. The radial pattern consists of concentric rings of tissue in the stem, hypocotyl and root of the seedling. The production of additional layers of cells results in an increase in size across the axis. The apical-basal axis has a characteristically polar nature, defined by functionally distinct structures rather than cell layers.

The importance of the apical-basal axis of development can be viewed in the evolutionary terms of its strong selective advantage in plant competition. Seedlings must rapidly reach the light through expansion along this axis whilst avoiding shading from competitors (Ballare, 1999). Out-competing its neighbours in terms of height promotes reproductive success through enhanced seed dispersal. The root meristem plays a crucial role in facilitating this success through the requisitioning of water and nutrient supplies. Thus the success of the polarised adult plant results from the correct establishment of the polar body plan during embryogenesis.

While embryology has long been an area of botanical investigation, the insights it can provide into fundamental aspects of developmental control, such as pattern formation, morphogenesis and differentiation have made it of critical importance to genetic and molecular studies.

1.1 How do we understand molecular mechanisms of development? Methods that have been used to analyse the genetic control of embryogenesis

Understanding the molecular mechanisms underlying embryogenesis can provide insight into developmental and metabolic regulation and the signalling systems integrating these processes. A great deal of research has been invested into analysing the genetic control mechanisms, exploiting a range of techniques to isolate genes of importance. The construction and screening of cDNA libraries from isolated RNA (Goldberg *et al.*, 1989) and promoter/enhancer trapping (Topping *et al.*, 1994) are just two of the techniques employed. A great deal of success has been achieved through mutational screens, highlighting genes that produce a knockout phenotype in the seed (Meinke and Sussex, 1979; Mayer *et al.*, 1991). Following the completion of the sequencing of the *Arabidopsis thaliana* genome (Arabidopsis Genome Initiative, 2000) research has focused on functionally characterising the 27,000 genes predicted (Ausubel and Benfey, 2002; Wortman *et al.*, 2003). Mutational studies have continued to play a major role in this analysis (Parinov *et al.*, 1999; Sessions *et al.*, 2002; Alonso *et al.*, 2003). Mayer *et al.* (1991) estimated that approximately 4000 genes were required for embryogenesis, with about 40 of these essential for pattern formation (Mayer *et al.*, 1991). Insertional mutagenesis screens have since demonstrated the requirement for a larger number of essential genes based on the resulting frequency of embryonic lethality (Franzmann *et al.*, 1995; McElver *et al.*, 2001; Tzafrir *et al.*, 2004). A number of different mutational strategies have been employed to highlight genes of interest and these are covered in the following section.

1.1.1 Mutational approaches

The generation and characterisation of genetic mutants with specific developmental defects, resulting from a single genetic lesion, is a relatively longstanding approach to uncover the control of development. These techniques can be employed in either a forward or reverse genetic approach.

The forward approach involves the screening of a mutagenised population for a phenotypic aberration of interest, for example one in which the apical or basal part of the embryo does not undergo development correctly. This approach can reveal genes of importance in the development of the structure of interest. However it is possible that the phenotype could result from a gene that is only indirectly related to the process.

The reverse genetics approach begins with the selection of a gene of interest. A mutagenised population can then be screened for mutants of the gene or a mutant directly created using gene expression masking techniques such as anti-sense RNA or RNAi. This method is generally preferable particularly in these post *Arabidopsis* genome sequence days, avoiding as it does the time consuming forward genetics approach to the cloning of mutant genes (Parinov *et al.*, 1999; Sessions *et al.*, 2002).

Mutant lines can be generated using one of three main methods: chemical, physical (irradiation) or insertional mutagenesis.

1.1.1.1 Chemical and physical mutagenesis

Chemical mutagenesis has been a major method employed in the creation of populations of novel mutants owing to its rapidity and the stable lesions formed (Topping and Lindsey, 1995). An advantage of both chemical and physical strategies is the production of large numbers of mutagenised plants, which can be screened for phenotypic abnormalities of choice. Most commonly this has consisted of soaking *Arabidopsis* seed in a solution of ethyl methane sulphonate (EMS). This powerful mutagen acts by reacting with guanine

residues, resulting in their alkylation. During subsequent DNA replication the alkylated guanine can pair with thymine rather than cytosine resulting in a transition type base substitution (Meinke & Sussex, 1979; King and Stansfield, 1997). A physical mutagenic strategy generally refers to the irradiation of seed using X-rays or gamma rays (Raghavan, 1997).

The generation of point or deletion mutations by these methods has allowed multiple alleles of a given mutant gene to be identified, thus allowing a comparison of phenotypic severity between mutant lines to deduce gene regions most critical for function (Topping and Lindsey, 1995).

A common difficulty of the approach comes in the form of homozygous lethal mutants, or those that are masked by gene redundancy. As previously alluded to, one of the major difficulties arising from these methods is that of cloning the disrupted gene. Advances such as CAPS marker analysis, which highlights an area of a chromosome in which the lesion has occurred, have accelerated the process (Baumbusch *et al.,* 2001).

1.1.1.2 Insertional mutagenesis

Insertional mutagenesis, a technique also known as gene tagging, provides a mutagenic approach from which cloning can be more effectively achieved. This approach involves the ectopic integration of a known DNA fragment into the plant genome. The consequent loss-of-function of the disrupted gene creates a mutant phenotype, while the insert also acts as a DNA tag to mark the site of integration, thus facilitating the recovery of the flanking plant DNA (unlike traditional chemical and irradiation mutagenesis techniques) (Topping and Lindsey, 1995).

The two major molecular tags, which have been employed in insertional mutagenesis approaches, have been transposons and the *Agrobacterium tumefaciens*-derived T-DNA system.

1.1.1.3 Transposon tagging

A transposable element is a sequence of DNA capable of moving around the genome. Mutant phenotypes can result from the insertion of the transposable element into a gene. Transposable elements were initially discovered in maize by McClintock (1950) and have since been isolated and characterised (reviewed in Pereira, 1998). Active transposable elements have been described in at least 35 plant species (Nevers *et al.*, 1986), including *Antirrhinum majus* (Coen *et al.*, 1989) and *Petunia hybrida* (Gerats *et al.*, 1990). *Arabidopsis thaliana* has only relatively recently been shown to have an active endogenous transposon system (Peleman *et al.*, 1991; Tsay *et al.*, 1993), prior to this heterologous transposons from other species have been used. The maize *Ac/Ds* system has been successfully introduced into *Arabidopsis* and shown to produce random, stable mutations (van Sluys *et al.*, 1987; Bancroft *et al.*, 1992; Dean *et al.*, 1992). The *Ac/Ds* system has been used to tag and clone a number of developmentally important genes in *Arabidopsis*, including *DRL1* (Bancroft *et al.*, 1993) and *PROLIFERA* (Springer *et al.*, 1995). A maize *En/I* system has also been used to tag the *MS2* gene (Aarts *et al.*, 1993).

An advantageous feature of the transposon tagging system is that excision of the transposable element may occur in later generations, thus providing confirmation of a direct linkage between mutant phenotype and the disrupted gene. However such excision events can leave a molecular footprint, which may complicate subsequent cloning.

1.1.1.4 T-DNA tagging

This approach has proved successful for much of the gene cloning and identification of recent years; its critical feature being the susceptibility of *Arabidopsis thaliana* to *Agrobacterium tumefaciens* mediated transformation.

Agrobacterium tumefaciens is a soil-borne plant pathogen capable of producing tumourous growths as a result of stable integration of bacterial DNA into the plant genome, and its subsequent expression (reviewed by Zupan and

Zambryski, 1995; Gheysen *et al.*, 1998). The requisitioning of the plants metabolism is due to the transfer of a sequence of DNA (referred to as transfer DNA or simply T-DNA), derived from the Ti-plasmid (tumour inducing) within the bacterium, into the plant genome (Chilton *et al.*, 1977; Schell *et al.*, 1979). The Ti-plasmid also harbours the *vir* (*virulence* region). This encodes the endonucleases necessary for T-DNA border cleavage and other proteins essential for mediating the transfer and integration of the T-DNA into the plant genome (Stachel and Nester, 1986). Secondary plant metabolites such as phenolic compounds produced by wounded plant tissue attract *Agrobacterium tumefaciens* (Hawes and Smith, 1989). Subsequent infection induces tumour formation at the wound sites characteristic of crown gall disease.

This naturally occurring system has been optimised for the integration of genetic material into the plant genome in a controlled manner (Bevan, 1984). Much of this optimisation involved the removal of plant hormone biosynthetic T-DNA genes (responsible for tumourous growth), whilst leaving the virulence sequences intact (reviewed in Gaudin *et al.*, 1994). Genes of interest can then be inserted in place of the removed oncogenes (Hoekema *et al.*, 1983; Zambryski *et al.*, 1983). In an insertional mutagenesis programme the T-DNA can be used as a tag to clone its flanking sequences and the gene that it has disrupted (Behringer and Medford, 1992).

Transgene integration is largely by illegitimate (non-homologous) recombination, and the random nature of this means that the entire genome is potentially susceptible to T-DNA integration (Gheysen *et al.*, 1991). Experimental evidence from studies using promoterless marker genes located in close proximity to a T-DNA border suggest that the majority of T-DNAs integrate into transcriptionally active regions (Koncz *et al.*, 1989; Topping *et al.*, 1991).

T-DNA insertional mutagenesis of *Arabidopsis thaliana* by *Agrobacterium tumefaciens* mediated transformation has proved a very popular means to link mutant phenotype to disrupted gene. Populations of transformants can be selected on antibiotic selection plates, before screening for mutant phenotypes of interest. Subsequent generations are subjected to segregation analysis to

ensure a link between the mutant phenotype and the T-DNA. Cloning of disrupted genes can be carried out using a number of techniques, including: Inverse PCR (IPCR); screening a genomic or cDNA library constructed from the mutant line; or by plasmid rescue (Behringer & Medford, 1992; Topping & Lindsey, 1995; Thomas, 1996).

A number of significant T-DNA insertion line collections have been established and made available to the research community, including those from the SALK Institute (Alonso *et al.*, 2003), Syngenta (Sessions *et al.*, 2002) and GABI-Kat (Rosso *et al.*, 2003). The Internet based services provided by the SALK Institute (http://signal.salk.edu/tabout.html) have facilitated a rapid access to mutated lines disrupted in genes of interest.

The study of plant development has encompassed a wide range of different species either as models for particular processes or those with commercial importance as crop plants. The majority of the work carried out into the molecular aspects of development in recent years has been centred around the study of the species *Arabidopsis thaliana,* which possesses a range of favourable characteristics elevating it to model plant status.

1.2 Arabidopsis as model species

Arabidopsis thaliana is a small cruciferous weed of the *Brassicaceae* (mustard) family, commonly known as thale cress, wall cress or mouse-ear cress (http://arabidopsis.org/info/aboutarabidopsis.jsp). It is in the same family as economically important crop species such as cabbage, broccoli and radish, yet is itself not of major agronomic importance (Meyerowitz, 1987). However, it has been used for studies in classical genetics for over sixty years, and a range of favourable characteristics have led to its widespread usage as a model species for the molecular and physiological study of plant development.

Arabidopsis plants consist of a rosette of small leaves and, following bolting, a main stem topped by an indeterminate floral meristem. At maturity the plant may reach a height of 30 or 40 cm. Its small size allows its cultivation in a

relatively limited space, in comparison to agricultural species such as soybean and maize, which may be hundreds of times larger (Walbot, 2000).

Flowers are typical of members of the mustard family, and normally self-fertilise (although cross pollination can be achieved by artificial manipulation). Each silique produced contains between 30 and 60 seeds, each less than a millimetre in diameter. As plants grown under favourable environmental conditions achieve a high seed set, a single plant may produce thousands of seeds.

Grown under long day conditions, with good nutrition, the commonly used ecotypes Columbia and Landsberg reach maturity in approximately six weeks (Meyerowitz, 1987). This relatively short generation time facilitates experimental procedures such as mutagenesis and genetic crosses, which require successive generations.

In addition to its developmental properties, which have made it such a convenient subject for classical genetic studies, *Arabidopsis* also has many advantageous genetic traits. *Arabidopsis* has one of the smallest known higher plant genomes, estimated to have a haploid size of approximately 125 million base pairs (Mb; The *Arabidopsis* Genome Initiative, 2000). This contrasts markedly with other plants used in genetic work such as tobacco, which has a haploid nuclear genome of 1600 Mb, and crop species such as wheat, which has a haploid genome of 5900 Mb (Meyerowitz, 1987).

The sequencing of the entire *Arabidopsis thaliana* genome has provided an incredibly powerful resource, allowing novel approaches to be utilised to rapidly identify and manipulate specific gene targets. Approximately 26,000 genes have been predicted from the sequence, though owing to 70% genome duplication, the actual figure is likely to be fewer than 15,000 different genes (Walbot, 2000). *Arabidopsis* genes are compact, containing several exons punctuated by short non-coding introns, and closely spaced, indicating short regulatory sequences. The *Arabidopsis* proteome is predicted to contain proteins from approximately 11,000 families (around 40% of which have more than five members), which is similar in functional diversity to that seen in the

nematode worm *Caenorhabditis elegans* and the fruitfly *Drosophila melanogaster*. 69% of the *Arabidopsis* genes were classified according to a sequence homology to proteins of known function in other organisms, though only 9% of these have been characterised experimentally. The approximately 30% remaining comprised those with plant-specific function and those with homology to genes of unknown function. While genes involved in such processes as vesicle trafficking and protein synthesis show homologies across the kingdoms, there are several classes of gene particularly abundant in *Arabidopsis* such as those involved in the synthesis or modification of cell walls. Proteins in families of more than five members appear highly numerous in *Arabidopsis* (The *Arabidopsis* Genome Initiative, 2000).

1.2.1 Genome Evolution

The completion of the *Arabidopsis* genome sequence revealed the extent to which the genome had been duplicated and consequently the degree of gene redundancy. *Arabidopsis* shares a common ancestor with the *Brassica* (cabbage and mustard) genera between 12 and 19 million years ago. There is still a high level of nucleotide conservation of their coding regions and their genomes are structurally similar. *Brassica* genomes have increased in size considerably through polyploidisation events, with extensive triplication and rearrangement. Analysis of the *Arabidopsis* genome suggests it has undergone two polyploidisation events. A duplication event 112 million years ago appears to have led to the generation of a triploid ancestor. Subsequent gene loss, additional local gene duplication events and chromosome rearrangement, have resulted in the small genome and five chromosomes seen today (The *Arabidopsis* Genome Initiative, 2000).

1.3 Embryogenesis

The many advantageous genetic traits of *Arabidopsis thaliana* have ensured its standing as a model plant for the study of plant genetics. Additionally it undergoes a well defined and highly reproducible embryogenesis, similar to

Capsella bursa-pastoris, a species long utilised as an anatomical model for embryogenesis. These factors combine to make *Arabidopsis* ideally suited to the study of the genetic control of embryogenesis.

1.3.1 Formation of the apical-basal axis of the embryo

The body plan of the plant shows a high degree of polarity, with its shoot-root axis. From the study of plant species such as *Arabidopsis* and *Capsella bursa-pastoris*, the first indication of the establishment of this polarity precedes even the fertilisation of the egg cell, which is present in alignment with the chalaza-micropyle axis of the asymmetrically organised embryo sac (Schultz and Jenson, 1968; Mansfield and Briarty, 1991). The egg cell, that following fertilisation gives rise to the embryo, is located within an embryo sac, which is itself surrounded by the maternal diploid tissue of the ovule. Polar organisation is evident in all of these structures along the apical-basal axis (Reviewed in Mansfield *et al.*, 1991; Gasser *et al.*, 1998; Schneitz *et al.*, 1998). Polar organisation within the embryo sac comprises the egg cell and synergids located at the micropylar end, while the antipodal cells occupy the chalazal end. The micropyle pole of the egg cell harbours a large vacuole in contrast to the chalazal pole that is relatively cytoplasmic (Schultz and Jenson, 1968). A double fertilisation occurs when the pollen tube enters the micropyle of the ovule delivering two haploid nuclei, one of which fuses with the egg cell nucleus (giving rise to the embryo) and the other with the central cell (giving rise to the endosperm). The double fertilisation has been implicated in the epigenetic control of gene expression, through the creation of differential gene dosage between embryo and endosperm (Lopes and Larkins, 1993). More dynamic examples of zygote polarisation are exhibited by *Papaver nudicaule* and maize, where fertilisation of the egg cell results in the movement of the cytoplasm and nucleus to the apical pole, exchanging positions with the vacuole (Olson and Cass, 1981; van Lammeren, 1981; Mòl *et al.*, 1994).

The marine intertidal brown algae (*Phaeophyceae*) contain the species *Fucus* and *Silvetia (*formerly *Pelvetia)*, which, due to some beneficial experimental features, have provided a model system in the study of the initial events of

zygote polarisation (Brownlee and Berger, 1995; Kropf, 1997; Brownlee and Bouget, 1998; Kropf *et al.*, 1999). Adult plants of both *Fucus* and *Silvetia* can be induced under laboratory conditions to release gametes. These relatively large free-living egg cells and zygotes are amenable to techniques such as microinjection and electrophysiology, which higher plant embryos and zygotes generally are not, in addition to fulfilling the requirement for large numbers of synchronously developing zygotes necessary for such studies (Brownlee, 2004). The zygote is produced by fertilisation away from the parent plant and is therefore under no direct influence from maternal tissue, and thus is presumed to have no inherent asymmetry. Fucoid zygotes can be polarised by external stimuli such as the direction of unilateral light (Jaffe, 1958) and fertilisation (Hable and Kropf, 2000). Following a polarizing event the axis formed is initially labile but becomes increasingly fixed (Brownlee, 2004). A number of physiological changes are associated with the axis formation including the formation of a cortical actin patch (in yeast and animal systems actin has a well established role in polarisation (Goode *et al.,* 2000)) at the rhizodermis within three hours of fertilisation; the establishment of an asymmetric distribution of mRNA and actin RNA (actin mRNA accumulates at the thallus pole during axis fixation) (Bouget *et al.,* 1996); polarised secretion and membrane redistribution (Shaw and Quatrano, 1996; Quatrano and Shaw, 1997; Belanger and Quatrano, 2000a, b); and the secretion of Golgi-derived cell wall components at the presumptive rhizoid pole (Kropf, 1997).

Studies in *Fucus* have provided a great deal of insight into the polarisation response to external stimuli, transposing that information from a free-living zygote to higher plants still leaves open the question as to when the apical-basal axis is first established. As already alluded to, the embryonic axis is aligned with that of the ovule, thus suggesting an orienting influence is derived from the maternal tissue. Indeed maternal mutations have been identified in *Drosophila*, which affect axis formation (Ingham, 1988; St Johnston and Nüsslein-Volhard, 1992). While such an influence is certainly suggested by the shared axis alignment, it is not a requirement and apical-basal pattern formation can be uncoupled from it, as demonstrated by apical-basal axis acquisition in somatic embryos in the absence of maternal cues (Nomura and Komamine, 1985; Luo and Koop, 1997). Additionally *in vitro* fusion of egg and sperm cell

has been shown in maize to form zygotes that acquire an axis of polarity prior to an asymmetric division event (Breton *et al.*, 1995).

Another extra-embryonic tissue, which might be assumed to play some role in the axis formation of higher plants, is the triploid endosperm, which arises from a fusion between a sperm cell and the diploid central cell of the ovule. Communication between embryo and endosperm has been shown to be essential for reproductive success, a situation highlighted by endosperm mutants such as the *miniature-1* mutant of maize that results in defective seed (Miller and Chourey, 1992). Despite its non-persistence (it degrades before embryo maturation) the endosperm plays numerous roles in embryogenesis including a role in embryo nutrition (Hirner *et al.*, 1998) and regulation of embryo size (Hong *et al.*, 1996). However, no conclusive evidence has as yet linked it to an involvement in embryo patterning.

The initial zygotic division in *Arabidopsis* and *Fucus* is asymmetric and transverse, generating two daughter cells of unequal size and fate. In *Fucus* the division leads to the generation of a larger upper cell that forms the thallus cell, and a smaller basal cell that forms the rhizoid and undergoes polarized growth. The differential secretion of cell wall components has been suggested to play a role in the determination of subsequent thallus and rhizoid cell identity. Wall fragments (sulphonated polysaccharides) have been shown to be sufficient to specify the fate of protoplasts of both cell types (Berger *et al.*, 1994), thus providing a foundation for the molecular basis of polarity in the *Fucus* zygote. Alternatively in the green alga *Volvox*, the specification of cell fate after the asymmetric division into vegetative and generative cells is attributed directly to the cell size rather than any unequal distribution of molecules (Kirk *et al.*, 1993). In *Arabidopsis* the apical cell is small, cytoplasmically rich and the site of highly active protein synthesis, while the basal cell derives from the vacuolar region of the zygote and is the larger of the two. The apical cell then divides into eight proembryo cells by two rounds of vertical and one round of horizontal division leading to the formation of the embryo proper.

By contrast, the basal cell undergoes repeated horizontal division to form a single file of between six and nine cells, the suspensor. These cells are all

initially extra-embryonic, although the uppermost cell of the suspensor, the hypophysis, adopts an embryonic fate and gives rise to part of the embryonic root meristem (Dolan *et al.*, 1993; Scheres *et al.*, 1994). The suspensor functions both as a conduit for nutrients to the developing embryo (Yeung and Sussex, 1979) and to position it in the lumen of the embryo sac (Yeung and Meinke, 1993).

In addition to size difference, the differing cell fate of the two daughter cells (arising from the initial zygotic division) is highlighted by the differential expression of the homeobox gene *Arabidopsis thaliana MERISTEM LAYER 1 (AtML1)*, whose gene transcript has been shown to accumulate in the apical cell, but is not detected in the basal cell (Lu *et al.*, 1996). The *WUS-RELATED HOMEOBOX (WOX)* genes also demonstrate differential expression dynamics revealing early embryonic patterning events. *WOX2* and *WOX8* are initially co-expressed in the egg cell and zygote but following the asymmetrical division of the zygote, *WOX2* and *WOX8* expression becomes confined to the apical and basal daughter cells respectively (Haecker *et al.*, 2004). Haecker *et al*, suggest that the *WOX2* and *WOX8* mRNA's are already arranged in a polar fashion within the zygote, and follow a similar course to the developmental determinants separated by asymmetric cell division in the *Caenorhabditis elegans* zygote (Lyczak *et al.*, 2002; Haecker *et al.*, 2004). Later in embryogenesis the expression patterns of these and other *WOX* genes become progressively restricted to smaller domains, defining distinct subpopulations of cells (Haecker *et al.*, 2004).

The clearest difference in fate is displayed by the non-permanent nature of the suspensor, with its programmed cell death occurring upon the embryo reaching its torpedo stage of development (Yeung and Meinke, 1993). Study of mutants showing abnormal cell division patterns of the suspensor provide clues as to the genetic control of suspensor cell identity. The *Arabidopsis* mutant, *abnormal suspensor* (Schwartz *et al.*, 1994) and the *raspberry* mutants (Yadegari *et al.*, 1994; Apuya *et al.*, 2002) arrest at the globular stage of embryogenesis, the suspensor subsequently takes on a range of characteristics normally associated with the embryo-proper including a series of inappropriate division. Expression of an apical marker *Arabidopsis thaliana LIPID TRANSFER*

PROTEIN (LTP) a homologue of the carrot *EP2* lipid transfer protein (Sterk *et al.,* 1991; Thoma *et al.,* 1994; Vroemen *et al.*, 1996) has been shown to occur in both the embryo-proper and the suspensor of *raspberry* (Yadegari, 1994).

The ability of suspensor cells to initiate embryogenesis in cases of embryo abortion and arrest gives rise to the suggestion that the embryo normally acts as a repressor of suspensor developmental potential, maintaining its differentiated state (Marsden and Meinke, 1985; Yeung and Meinke, 1993). In seeds of the *twin (twn)* mutant of *Arabidopsis* a re-differentiation of suspensor cells gives rise to a secondary embryo, alongside abnormal development of the primary embryo. The additional embryo has normal or opposite apical-basal polarity (Schwartz *et al.*, 1994; Vernon and Meinke, 1994; Zhang and Somerville, 1997), thus implicating the interaction between embryonic and extra-embryonic cells in the orientation of the embryos apical-basal axis. A MAPKK kinase gene, *YODA (YDA)*, has been identified that is required for correct partitioning of embryonic and extra-embryonic cell fates. In loss-of-function mutants zygote elongation is suppressed, and cells of the basal lineage take on an aberrant embryonic fate, rather than forming the extra-embryonic suspensor. In the converse gain-of-function mutant, proembryo development is suppressed and the suspensor undergoes exaggerated growth. At it's most severe the zygote develops into a file of cells with no detectable proembryo. This suggests that the asymmetric division results in a selective inactivation of *YDA* in the apical daughter cell, allowing the establishment of embryonic cell fate (Lukowitz *et al.*, 2004). A role has been suggested for auxin in the promotion of embryonic fate in the apical lineage, this is complementary to the role of *YDA* in the promotion of zygotic elongation and extra-embryonic fate in the basal lineage (Friml *et al.*, 2003; Lukowitz *et al.*, 2004). The signalling events required to initiate the MAP kinase cascade, and therefore control the determination of embryonic versus extra-embryonic cell, are unclear but Lukowitz *et al.* (2004) speculate that it could be an extra-cellular signal from the endosperm or the maternal seed coat.

So far a range of extra-embryonic factors have been reviewed in terms of their potential involvement in the orientation and formation of the apical-basal polarity of the embryo. In addition to these external factors, what is known about

embryonically expressed genes that may play a role in the establishment of polarity at zygotic division? The *Arabidopsis* mutant *fass* undergoes a symmetrical initial zygotic division, despite the resulting aberrant morphology; apical-basal polarity of the embryo is still established. Thus demonstrating that a stable asymmetry can be established despite the absence of an asymmetrical first zygotic division (Torres Ruiz and Jürgens, 1994; Fisher *et al.*, 1996). A systematic screen of embryo defective mutants has implicated only a single gene *GNOM/EMBRYO DEFECTIVE 30* (*GN/EMB30*) so far, that has been shown to interfere with the stable establishment of the apical-basal axis in the embryo (Mayer *et al.*, 1991, 1993). The earliest of the defects manifested in *gn* mutant embryos are at the zygotic stage of development. The *gn* zygote fails to elongate to such an extent as the wild-type, this is then followed by a variable (rather than asymmetric) division of the zygote, and subsequent abnormal cell division patterns (Mayer *et al.*, 1993). In addition to its role in axis formation, *GNOM* has been found to be generally required for the stable maintenance of plant cell polarity. Loss-of-function mutants of *GNOM* have severely altered phenotypes based upon the abnormal cell division events; these include lack of an embryonic root and the fusion (or deletion) of the hypocotyl and cotyledons. In its most severe form, the lack of apical-basal polarity results in ball-shaped seedlings with differentiated, but randomly oriented vascular cells in their centre (Mayer *et al.*, 1993; Steinmann *et al.*, 1999). In culture, *gn* mutant cells fail to acquire a common polarity and are thus incapable of organised growth. The lack of stable apical-basal axis fixation is evidenced in *gn* mutants by strong markers of apical and basal polarity, such as *AtLTP1* and *POLARIS* respectively, showing variable or no sign of correct apical-basal polarity (Vroemen *et al.*, 1996; Topping and Lindsey, 1997).

Many aspects of the *gn* mutant phenotype can be phenocopied *in vitro*, in cultured embryos of the closely related species *Brassica juncea*, that have been subjected to high concentrations of auxin analogues or auxin transport inhibitors (Liu *et al.*, 1993; Hadfi *et al.*, 1998). These observations would appear to suggest a role for auxin in the establishment of early embryo polarity, and create a central role for *GNOM* in polar auxin transport. The *GNOM* gene is expressed throughout development and encodes a guanine-nucleotide exchange factor (GEF) for small GTP-binding proteins of the ARF family

(Shevell *et al.*, 1994; Busch *et al.*, 1996; Steinmann *et al.*, 1999). It shows sequence similarity to two yeast proteins Gea1p (Peyroche *et al.*, 1996) and Gea2p (Busch *et al.*, 1996), GEFs involved in the transport of vesicles between the endoplasmic reticulum and the Golgi complex. The GNOM protein is susceptible to inhibition by the fungal toxin brefeldin A (BFA) (Steinmann *et al.*, 1999). This compound has also been shown to inhibit targeted wall secretion and the fixation of the polar axis in *Fucus* (Shaw and Quatrano, 1996), resulting in similarly defective zygotic division and apical-basal patterning as the *gn* mutant. Brefeldin A also affects the co-ordinated polar localisation of the putative auxin efflux carrier PIN FORMED 1 (PIN1), which is normally localised to basal side of transporting cells (Bennett *et al.*, 1996; Galweiler *et al.*, 1998; Palme and Galweiler, 1999; Geldner *et al.*, 2001; Friml *et al.*, 2003). The involvement of the *GNOM* gene in the localisation of PIN1, and consequently the maintenance of polarity was deduced using an engineered BFA-resistant *GNOM* variant (Geldner *et al.*, 2003). In transgenic BFA-resistant *gnom* mutant plants, PIN1 polar localisation and thus auxin transport was unaffected in the presence of BFA, suggesting that the role of *GNOM* in the establishment of the apical-basal axis is directed through the specific support of PIN1 localisation. However, the basal localisation of PIN1 does not begin until the 16- to 32- cells stage, thus requiring an alternative explanation for the prior breakdown of polarity. Directed vesicle transport in the establishment of cell polarity in mammalian cells would suggest that prior to its involvement in PIN1 localisation *GNOM* may be important in the vesicle targeting of other factors (Mostov *et al.*, 2000; Gruenberg, 2001; Geldner, 2004).

Friml *et al.* (2003), demonstrated a dynamic distribution of auxin through early embryogenesis, and postulated that these gradients were central to the establishment of the apical-basal axis. Auxin accumulates in the apical daughter cell following the initial zygotic division; transport from the basal cell is predicted to be PIN7-dependent. Supporting evidence for this putative auxin efflux carrier come from its polar localisation to the apical plasma membrane of the basal cell, and that *pin7* mutants fail to establish the apical-basal auxin gradient resulting in the accumulation of auxin in the basal cell. Auxin production is predicted to commence in the apical region of the embryo by the globular stage, concurrently PIN1 becomes basally localised in the provascular

cells, and PIN7 localisation is reversed within the basal cells. Combined with the initiation of *PIN4* expression at the basal pole of the embryo, these actions suggest auxin transport is being directed towards the future root pole and out of the embryo (Friml *et al.*, 2003). Recently, chemical and genetic inhibition studies have suggested that PIN1 and PIN4 proteins play a major buffering role to maintain these auxin gradients in the embryo (Weijers *et al.*, 2005). Interestingly the most severe *YDA* gain-of-function phenotype in many ways phenocopies the most severe *pin7* loss-of-function phenotype, this has been interpreted to confirm the independent operations of YDA promoting extra-embryonic fate in the basal lineage, and auxin promoting embryonic fate in the apical lineage (Lukowitz *et al.*, 2004).

1.3.2 Later stages of embryo development

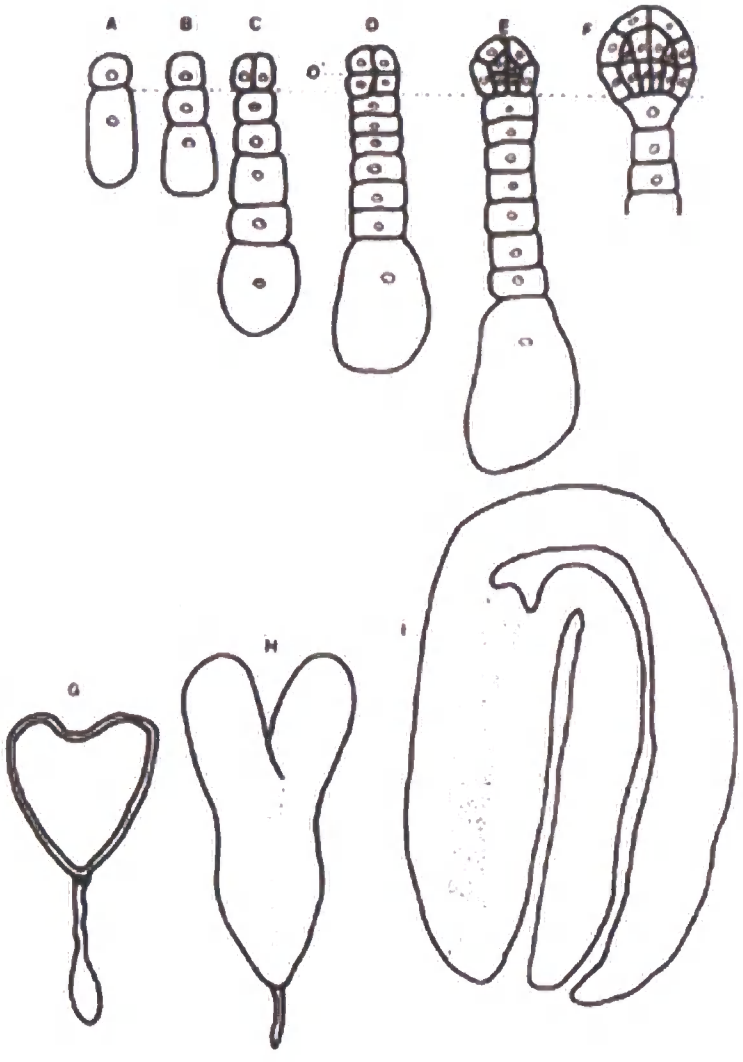
Embryogenesis in the plant kingdom is not a uniform process, and there is considerable difference in its morphological aspects even between relatively closely related species (Steeves and Sussex, 1989). The model plant species *Arabidopsis thaliana* undergoes a similar pattern of embryogenesis to *Capsella bursa-pastoris*, a species long utilised as an anatomical model for embryogenesis (see Figure 1.1) (Schultz and Jenson, 1968; Mansfield and Briarty, 1991). Embryogenesis in *Arabidopsis* is a continuous process, although for convenience it can be separated into three major phases: early, mid and late, as will be described below. The first phase is one of pattern formation and morphogenesis, during which the axes of the plant body plan are defined, and organ systems formed. The second phase is that of maturation, with a characteristic accumulation of storage reserves. In its final phase the embryo prepares for developmental arrest. *Arabidopsis* embryogenesis is rapid, with the early and mid phases completed eleven to twelve days post fertilisation, and only fourteen days to the completion of the late phase and the production of desiccated mature seed (see Figure 1.1) (Lindsey and Topping, 1993).

As previously described, the process of embryogenesis in *Arabidopsis* is initiated by a double fertilisation. This contrasts with the single fertilisation in

Figure 1.1 Schematic representation of embryogenesis in *Arabidopsis thaliana* and *Capsella bursa-pastoris*.

A – C	Proembryonic stages
D	Octant stage (O' indicates the position of the designated O' boundary)
E, F	Globular stage
G	Heart stage
H	Torpedo stage
I	Cotyledonary stage

The dotted line represents the distinction between the embryo proper and the extra-embryonic suspensor.



Taken from Lindsey and Topping (1993).

gymnosperms, which gives rise to a single diploid cell, which acts as the progenitor for both embryonic and storage reserve tissues (Raghavan, 1997).

Embryogenesis commences with an approximate three-fold elongation of the zygote, before a transverse, asymmetrical division generates apical and basal cells (Schultz and Jenson, 1968; Tykarska, 1976; Mansfield and Briarty, 1991). In dicotyledonous species, the apical and basal cells differentiate as previously described into the embryo proper and the suspensor. In monocotyledonous species, the apical cell forms the embryo proper in addition to the majority of the suspensor; the basal cell does not divide further and forms only the terminal cell of the suspensor.

Even at the first zygotic division plant development between species begins to diverge, with some plants exhibiting a symmetric division, which may be in an oblique or longitudinal orientation (Sivaramakrishna, 1978). The details of plant embryogenesis amongst different plant species vary considerably, leading to a common end point in the generation of a developmentally arrested mature embryo.

In *Arabidopsis* a series of highly reproducible symmetrical divisions forms a structure of eight isodiametric cells, referred to as the octant stage. In other species, such as maize (*Zea mays*), development is not so defined and reproducible, comprising a series of predominantly asymmetrical divisions after the initial zygotic division (Sheridan and Clarke, 1987). The eight cells of the *Arabidopsis* octant stage embryo are organised into two layers (upper and lower tier) of four cells, separated by the so-called O' boundary (Tykarska, 1976, 1979). The upper tier contributes to the cotyledons and the shoot apex, while the lower tier forms the hypocotyl and the majority of the embryonic root. Sector analysis suggests that the lower tier may also contribute to the basal region of the cotyledons (Scheres *et al.*, 1994). Proceeding from the octant stage, an anticlinal division (at right angles to the outer surface) gives rise to the eight-cell protoderm (outer layer) of the sixteen-cell embryo. Continued anticlinal division in the protoderm cells coupled with longitudinal and, later, transverse divisions of the interior cells gives rise to the radially symmetrical globular stage embryo, approximately two or three days post-fertilisation. The

globular stage is concluded when the embryo comprises approximately thirty cells and measures approximately 40 μm in diameter (Lindsey and Topping, 1993). There is a polar distribution of shoot apical meristem (SAM) determinants in the upper region of the globular embryo, with centred expression of essential SAM genes, although no morphologically recognisable structure is yet in place (Barton and Poethig, 1993; Long *et al.*, 1996; Mayer *et al.*, 1998). Clonal analyses in cotton and maize suggest that the embryonic region destined to form the shoot apical meristem is determined in the early globular stage (Christianson, 1986; Poethig *et al.*, 1986).

During the transition stage, differential rates of cell division between the centrally located epiphyseal cells and the cotyledon progenitor cells lead to the formation of a bilaterally symmetrical heart stage embryo, with its distal lobes representing the precursors of the cotyledons (Swamy and Krishnamurthy, 1977). Fate mapping has shown the generation of these cotyledon primordia to be a sequential series of events (Woodrick *et al.*, 2000). Approximately three or four days post fertilisation the embryo now comprises approximately two hundred and fifty cells (Lindsey and Topping, 1993). By this stage the three basic tissues (epidermis, ground tissue and vascular tissue) of the embryo have been established in rudimentary form, marking the end of early embryogenesis, with morphological pattern and polarity established (Mayer *et al.*, 1991; Jürgens *et al.*, 1994).

The heart stage embryo of *Arabidopsis* represents the basic unit of the plant, although further refinement of the embryonic pattern occurs during subsequent developmental stages. Mid-embryogenesis is characterised by rapid cell division to define the cotyledons and a later cell enlargement, especially in the lower region to enlarge the hypocotyl. The embryo also begins to prepare itself for late embryogenesis by a simultaneous synthesis of protein, oil and starch reserves, which accumulate in the cotyledons and the endosperm (Pang *et al.*, 1988). In a situation unique to seed plants, morphogenesis is interrupted in the later stages of maturation by the accumulation of storage reserve macromolecules in virtually all cells of the embryo, playing a large part in the rapid increase in embryo size and mass (Walbot, 1978; Tykarska, 1987; West and Harada, 1993). In post-heart stage embryos a SAM becomes discernable

as three distinct cell layers, although it is quite undeveloped and remains so until after germination (Barton and Poethig, 1993). The suspensor begins to senesce during the torpedo stage and is completely absent by the mid-cotyledonary stage. The endosperm also starts to be consumed until at maturity it comprises only a tiny proportion of the seed, in direct contrast to graminaceous species, whose seeds are highly endospermic. There is a characteristic folding back of the cotyledons at the bent cotyledonary stage where the embryo comprises approximately twenty thousand cells and is approximately 500 µm in length despite its folded state (Lindsey and Topping, 1993). By this stage the hypocotyl and cotyledons are very well defined and the patterning of the embryonic root meristem is clearly established (Scheres *et al.*, 1994).

During the final phase (late embryogenesis) the embryo becomes metabolically quiescent, dormant and desiccation tolerant, with a 10-90% loss of water prior to entering dormancy (Goldberg *et al.*, 1989; Kermode, 1990). The synthesis of LEA (Late Embryogenesis Abundant) proteins is thought to be in a protective capacity during the desiccation step (Dure, 1993).

The embryo gradually develops the competence to germinate. Abscisic acid (ABA) suppresses precocious germination and imposes dormancy. This is broken by the growth regulator gibberellin (GA), which promotes germination through a poorly understood mechanism (Ritchie and Gilroy, 1998). Mutants of the *COMATOSE* (*CTS*) locus do not respond to gibberellin and display a marked reduction in germination potential. It is postulated that *CTS* may be required to remove embryo dormancy to allow germination (Russell *et al.*, 2000). Mutants of the *Arabidopsis ABSCISIC ACID-INSENSITIVE3* (*ABI3*) and the maize *VIVIPAROUS-1* (*VP1*) genes display a reduced sensitivity to exogenous ABA and fail to become dormant. In addition they contain significantly reduced levels of storage protein. It has been proposed that ABI3 and VP1 proteins could act as transcriptional activators, interacting with transcription factors from various pathways to control gene expression during seed development (Parcy *et al.*, 1994). The *leafy cotyledon* (*lec1*; Meinke, 1992; Goldberg and Harada, 1994; Meinke *et al.*, 1994; West *et al.*, 1994; Lotan *et al.*, 1998) and *fusca3* (*fus3*; Bäumlein *et al.*, 1994; Keith *et al.*, 1994; Reidt *et*

al., 2000; Tsuchiya *et al.*, 2004) mutants have unaltered ABA sensitivity but display *ABI3* phenotypic characteristics, such as a reduced storage protein accumulation, a failure to become dormant, and desiccation intolerance. *ABI3*, *FUS3* and *LEC1* have been demonstrated to act synergistically to regulate many elementary processes involved in mid and late embryogenesis (Parcy *et al.*, 1997). The cotyledons of *lec1* and *fus3* mutant embryos display leaf-like characteristics such as trichomes, and a cellular organisation intermediate between that of wild type cotyledons and leaves (Meinke *et al.*, 1992; Keith *et al.*, 1994). Another gene, *BABY BOOM*, encodes an *AP2* domain transcription factor, which activates signal transduction pathways for the induction of embryonic development from differentiated somatic cells (Boutilier *et al.*, 2002). In the *pickle* (*pkl*) mutant of *Arabidopsis*, embryonic differentiation characteristics are present after germination, this is thought to be due to *PKL* acting as part of a gibberellin modulated developmental switch to prevent retention of the embryonic developmental state (Ogas *et al.*, 1997; Ogas *et al.*, 1999). These mutants are helpful in elucidating the function of these key regulators of late embryonic development.

1.3.3 Mechanisms for the control of pattern formation

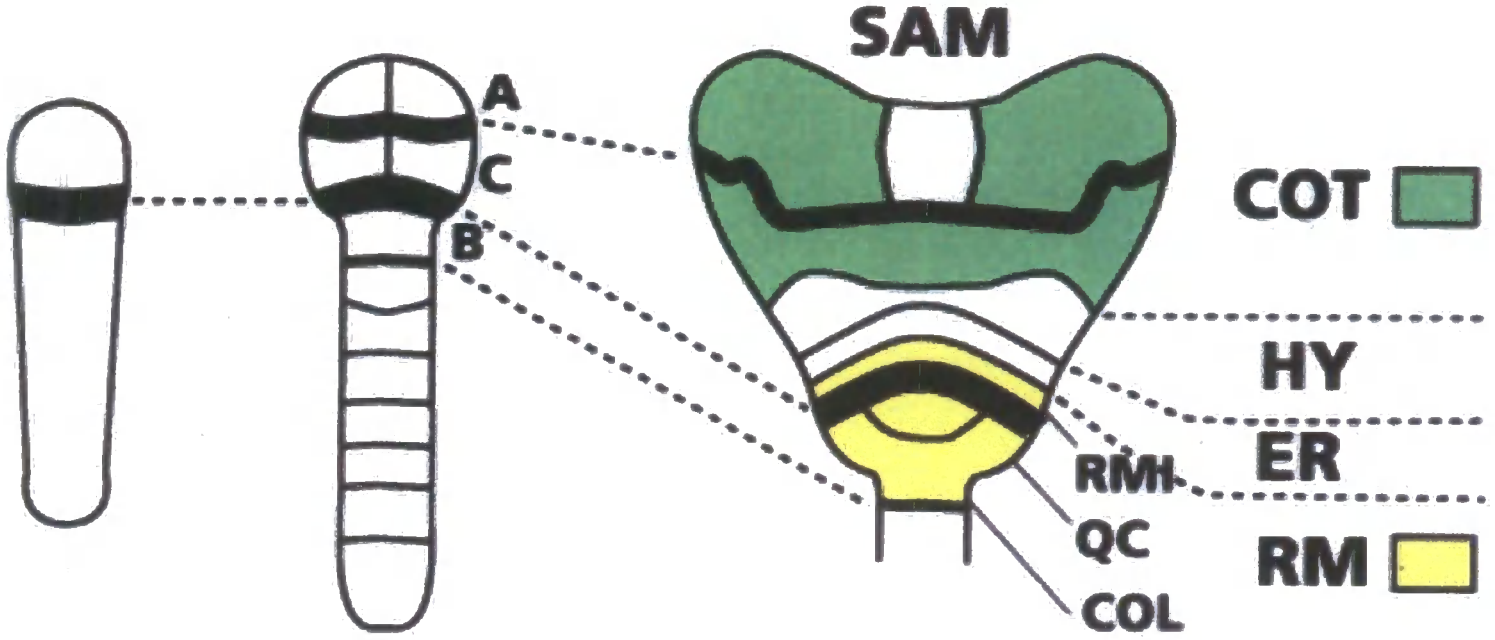
The apical-basal pattern of the *Arabidopsis* embryo is defined by the positioning of the shoot meristem and cotyledons, the hypocotyl and the root and root meristem. This pattern is the result of highly reproducible cell divisions through development, such that the resultant seedling structures can be traced back to their progenitors in the early stages of embryogenesis (Tykarska, 1976, 1979; Mansfield and Briarty, 1991). The study of mutant phenotypes of *Arabidopsis* embryogenesis has led to the theory that the embryonic axis is initially partitioned into three main regions: apical, central and basal (see Figure 1.2) (Mayer *et al.*, 1991). It should be noted that, while these regions do display differing developmental commitments, they should not be considered as entirely independent in terms of regulation, as clear interactions between tissues has been shown to be essential for correct global patterning of the seedling (Laux and Jürgens, 1997). The apical region gives rise to the shoot meristem and the majority of the cotyledons, although the central region contributes to their

Figure 1.2 Clonal regions of the embryo

The study of mutant phenotypes of *Arabidopsis* embryogenesis has led to the theory that the embryonic axis is initially partitioned into three main regions: apical, central and basal (Mayer *et al.*, 1991)

A = Apical
C = Central
B = Basal

SAM = shoot apical meristem
COT = cotyledons
HY = hypocotyl
ER = embryonic root
RMI = root meristem initials
QC = quiescent centre
COL = columella initials



(Figure taken from http://www.bio.uu.nl/mg/pd/research/index.html)

shoulder region (Scheres *et al.*, 1994). In addition the central region contributes the majority of the remaining axis, the hypocotyl, the vascular cortex, the embryonic root and root meristem endodermal initials. The apical and central regions derive from the upper and lower tiers respectively of the octant stage embryo and therefore by extension the apical cell daughter cell arising from the initial zygotic division. The basal region, which gives rise to the quiescent centre, the columella initials and the central root cap is clonally separate corresponding to the descendents of the hypophyseal cell, the uppermost cell of the suspensor, a derivative of the basal zygotic daughter cell (Scheres *et al.*, 1994; Mayer and Jürgens, 1998). All three regions are established by the octant stage with the formation of the O' boundary portioning the upper and lower tiers. Cells in the apical region divide without preferential orientation, while divisions perpendicular to the axis create the cell files of the central region. The generation of the root meristem and central root cap require a stereotyped set of divisions, such that the fate of cells in this region can be predicted with a high degree of probability (Scheres *et al.*, 1994). The position of a cell along the embryonic apical-basal axis appears to determine its fate, this seems to be the case not only plant species such as *Arabidopsis* but also in those with less reproducible patterns of cell division. In order to generate such a tightly regulated system of gene expression in a position related, region specific manner, there must be control mechanisms in place. In the following sections, genes involved in the specification of cell fate along the apical-basal axis will be discussed along with possible signalling control mechanisms.

1.3.3.1 The apical region

The primary shoot meristem of the adult plant is a stem cell system providing the ultimate source for all the plants aerial parts; lateral primordia are initiated at its flanking regions in a specific pattern termed phyllotaxis (Fleming, 2005). As previously described, the apical region of the embryo forms the self-perpetuating shoot meristem and the majority of the cotyledons. It was initially believed that the cotyledons were in fact the first products arising from the shoot meristem, however, while meristematic gene expression is detectable in the embryo from an early stage, not all the required components of the shoot

meristem are present by the heart stage at which point the presumptive cotyledons have already started to emerge (Kaplan and Cooke, 1997; Souter and Lindsey, 2001). In the mature embryo the shoot meristem comprises approximately one hundred cells, this precise size restriction is essential for the shape of the plant (Irish and Sussex, 1992). A number of genes have been identified which have an effect on the establishment or maintenance of the shoot meristem (Laux and Mayer, 1998).

An early marker of shoot meristem fate is the *WUSCHEL* (*WUS*) gene, which encodes a novel homeodomain transcription factor; this is first detected in the four inner apical cells at the sixteen-cell stage of *Arabidopsis* embryogenesis (Mayer *et al.*, 1998). *WUS* expression is maintained through a number of asymmetrical divisions in a group of cells, which at this stage also include some that go on to produce the cotyledon primordia. By the globular stage, expression is restricted to a group of cells (L3 layer cells) beneath the subepidermal cells (L2 layer cells) at the apex of the embryo, which will become the shoot meristem (Laux *et al.*, 1996). These cells are proposed to act as an organising centre through which *WUS* acts to maintain the pluripotent capacity of the slowly dividing stem cells above (Mayer *et al.*, 1998; Lenhard and Laux, 1999). The mechanism by which the cells below this putative organising centre are not induced to adopt stem cell identity is as yet unclear (Torres Ruiz, 2004).

The *CLAVATA* signalling pathway has emerged as a negatively regulated feedback loop required for controlling the size of the stem cell population (Schoof *et al.*, 2000). The expression of *CLAVATA 1 (CLV1),* a membrane-bound receptor serine/threonine kinase is initiated in the stem cell population of the heart stage embryo (Clark *et al.*, 1993, 1996). The expression of its ligand *CLAVATA 3* (*CLV3*) is activated in the shoot meristem stem cells in respond to the influence of *WUS* from the organising centre (Clark *et al.*, 1995; Fletcher *et al.*, 1999; Trotochaud *et al.*, 2000). *CLV3* binds to *CLV1* at the plasma membrane, activating *CLV1* and causing repression of the *WUS* gene at the transcript level (Brand *et al.*, 2000; Waites and Simon, 2000; Clark, 2001). The *CLV* genes promote the transition to organogenesis; *POLTERGEIST* has been implicated as a possible downstream target of *CLV3* as part of this process (Yu *et al.*, 2000).

Cells in the central region of the embryo axis express *SHOOT MERISTEMLESS* (*STM*), a *KNOTTED1*-like homeobox *(KNOX)* gene, from the late globular stage (Long *et al.*, 1996). *stm* mutant embryos are blocked in the development of a shoot meristem from the torpedo stage of embryogenesis onwards, but are otherwise normal (Barton and Poethig, 1993). Tobacco plants constitutively expressing the maize *KNOTTED1* gene develop many ectopic shoot meristems, demonstrating the converse gain-of-function phenotype (Sinha *et al.*, 1993). SAMs observed in weak mutant alleles of *stm* cease organ initiation after the production of only limited numbers of leaves and floral organs, thus there appears to be a requirement for *STM* in both SAM formation and the maintenance of post-embryonic SAMs (Clark *et al.*, 1996). Expression of *STM* is switched off at the organ primordia initiation sites of cotyledons and leaves, expressed in its place is a myb-domain transcription factor *ASYMMETRIC LEAVES 1* (*AS1*) (Byrne *et al.*, 2000). The partial rescue of the *stm* single mutant phenotype in *stm as1* double mutants indicates that the role of *STM* in maintaining the undifferentiated state of cells with a shoot meristem fate is mediated by the repression of the primordia promoting *AS1* gene. In the seedling both *AS1* and *AS2* act to maintain leaf identity through the repression of class I *KNOX* genes (Byrne *et al.*, 2000; Semiarti *et al.*, 2001; Iwakawa *et al.*, 2002).

The *ROUGH SHEATH2 (RS2)* gene of maize and the *PHANTASTICA (PHAN)* gene of *Antirrhinum* both encode proteins homologous to *AS1* and act to negatively regulate class I *KNOX* genes (Waites *et al.*, 1998; Timmermans *et al.*, 1999; Tsiantis *et al.*, 1999). *AS2* is a novel nuclear protein expressed in the adaxial domain of the cotyledons (Iwakawa *et al.*, 2002). It is proposed that the *PICKLE (PKL)* and *SERRATE (SE)* genes control the expression of *AS1* and *AS2* via a regulation of chromatin structure (Ogas *et al.*, 1999; Ori *et al.*, 2000; Eshed *et al.*, 2001; Prigge and Wagner, 2001). *AS1*, *AS2*, *PKL* and *SE* act in concert to maintain the differentiated state of leaf cells through the negative regulation of the genes that lead to the formation of meristematic tissues.

The observation that *CLV1* and *WUS* expression is initiated but not maintained in *stm* mutants, and that *STM* expression is not maintained in *wus* mutants,

suggest that while *WUS* and *STM* are established independently they are mutually dependent for subsequent expression (Long and Barton, 1998; Mayer *et al.*, 1998).

Another gene that interacts with *STM* is *ZWILLE* (*ZLL*), a member of a novel gene family implicated in translational control (Moussian *et al.*, 1998; Lynn *et al.*, 1999). *zll* mutants correctly initiate shoot meristems, but display a restricted or down-regulated expression of *STM*, consequently cells follow non-meristematic fate. *ZLL* therefore appears to be required for maintenance of meristematic cell identity within the apex. After the establishment of the leaf primordia the meristem appears able to regulate itself without a requirement for *ZLL* (Moussian *et al.*, 1998).

Sites of primordia initiation are uncovered by specific gene expression patterns. Cells destined for floral meristem initiation express a putative transcription factor, *REVOLUTA* (*REV*), known to be involved in the formation of secondary meristems (Talbert *et al.*, 1995; Otsuga *et al.*, 2001). *AINTEGUMENTA* (*ANT*), a transcription factor involved in the maintenance of an organs proliferative state, indicates an organ primordia initiation site (Elliot *et al.*, 1996; Mizukami and Fischer, 2000). The generation of phyllotaxis is based upon the position of new primordia initiation sites being influenced by existing primordia. Control mechanisms for this phenomenon are covered later.

The complex network of interactions that set up the shoot apex originates from the apical region of the globular stage embryo. This region is partitioned into two expression domains. Initially *ANT* is expressed in a ring-shaped peripheral domain, with concentrated expression at the sites of presumptive cotyledon formation (Elliot *et al.*, 1996). *STM* expression is activated in a cell of the peripheral domain. Its expression then spreads in a stripe across the central region to partially overlap the ring shaped peripheral domain (Long and Barton, 1998). Three partially functionally redundant NAC domain transcription factors, *CUP-SHAPED COTYLEDON1* (*CUC1*), *CUC2* and *CUC3*, act to suppress fusion of the cotyledon primordia, and regulate the formation of the embryonic SAM by the induction of *STM* expression (Aida *et al.*, 1997, 1999; Takada *et al.*, 2001; Vroemen *et al.*, 2003). The *Arabidopsis* CUC proteins are

homologous to the *Petunia* NO APICAL MERISTEM (NAM) and *Antirrhinum* CUPULIFORMIS proteins (Souer *et al.*, 1996; Weir *et al.*, 2004).

CUC2 exhibits a similar expression pattern to that of *STM* during the globular and heart stages of embryogenesis (Aida *et al.*, 1997). During the walking-stick stage the expression patterns separate such that *STM* is restricted to the centrally located primary shoot meristem, while *CUC2* defines a peripheral region between the cotyledon bases (Aida *et al.*, 1997, 1999). *STM* expression is completely blocked in the *cuc1 cuc2* double mutant demonstrating its dependence on *CUC1* and *CUC2* (Aida *et al.*, 1997, 1999, 2002). *CUC1* and *CUC2* mRNAs are detectable in *stm* even during late embryogenesis, but are disturbed, suggesting the requirement for *STM* in the precise patterning of CUC expression (Takada *et al.*, 2001). *CUC* genes are expressed between cotyledon (and floral) primordia suggesting that they function to prevent fusion through the restriction of boundary region growth between organ primordia (Ishida, 2000; Takada, 2001). The fusion of cotyledons and floral organs observed in *stm*, and the location of *STM* expression at organ primordia boundaries also suggest the involvement of STM in organ separation (Long and Barton, 2000). The astounding complexity of the apical pole organisation becomes apparent when such micro networks of expression interaction are linked together in a spatial and temporal framework.

1.3.3.1.1 Patterning the cotyledon region

The polar localisation of gene expression as described above indicates a clear separation of cell fate domains for the shoot meristem versus the cotyledon primordia, from an early stage of embryogenesis. The cells of the cotyledon primordia have a distinct expression fingerprint distinguishing them for instance from cells initiating adult leaf primordia. Cotyledon primordia cells are not derived from those that have previously expressed *STM* (Long *et al.*, 1996).

Cotyledons display many properties of rosette leaves, but lack or gain specific features. However, observation of the mutant phenotypes of certain genes involved in late embryogenesis such as *LEC1* and *FUS3* show cotyledons can

be partially converted to leaf-like organs (Meinke, 1992; Lotan *et al.*, 1998; Stone *et al.*, 2001). The converse cotyledon-like leaf phenotype has also been described indicating that they are homologous structures (Conway and Poethig, 1997).

The transition stage embryo acquires bilateral symmetry through the establishment of the two cotyledon primordia. A sequential establishment of the two primordia has been demonstrated, in keeping with the spiral phyllotaxis model (Woodrick *et al.*, 2000). The initial leaf primordia are established approximately at right angles to the cotyledons but with subsequent primordia the lateral inhibition effect brings the angle down to 137.5° (Reinhardt, 2005).

Loss-of-function mutants in the *GURKE*, *PASTICCINO* and *PEPINO* genes lead to impaired cotyledon formation (Torres Ruiz *et al.*, 1996; Faure *et al.*, 1998; Vittorioso *et al.*, 1998). The mutant phenotypes displayed are quite pleiotropic but seem to control proliferation particularly of aerial structures. Cloning has placed *PASTICCINO* and *PEPINO* in a highly conserved pathway, but there is presently no molecular model for early organisation of the apical domain (Vittorioso *et al.*, 1998; Bellec *et al.*, 2002; Haberer *et al.*, 2002).

Despite any similarities shared with adult leaves, cotyledon structure is not particularly impaired in mutants of *ANT, ZLL, REV, AS1* or *AS2* (Elliot *et al.*, 1996; Moussian *et al.*, 1998; Lynn *et al.*, 1999; Byrne *et al.*, 2000; Mizukami and Fischer, 2000; Otsuga *et al.*, 2001; Semiarti *et al.*, 2001 Iwakawa *et al.*, 2002). However, their correct adaxial or abaxial localisation of expression pattern in the cotyledons appears to be essential for normal SAM development.

SAMs are formed on the adaxial side of leaves, and a reciprocal signalling relationship between adaxial fate and SAM formation has long been suggested by classical experiments (Sussex, 1954; Snow and Snow, 1959). Takada *et al.* (2001) demonstrated that ectopic expression of *CUC1* led to the formation of adventitious SAMs on the adaxial side of the cotyledons, suggesting that only the adaxial side has the competence to form meristems.

The *ZWILLE/PINHEAD (ZLL/PNH)* gene is adaxially expressed and functions redundantly with the related gene *ARGONAUTE (AGO)* to regulate leaf polarity and SAM formation (Bohmert *et al.*, 1998; Moussian *et al.*, 1998; Lynn *et al.*, 1999). Mutations in *ZLL/PNH* lead to formation of a terminal leaf-like organ from the arrested SAM (Moussian *et al.*, 1998; Lynn *et al.*, 1999). A double mutant in *ZLL/PNH* and *AGO* fails to express the *STM* gene, this suggested they are required for SAM formation (Lynn *et al.*, 1999).

The promotion of SAM formation through adaxial identity is also demonstrated by gain-of-function mutations in *PHABULOSA (PHB)* and *PHAVOLUTA (PHV)*, which lead to adaxialisation of the cotyledons, leaves and floral organs and the formation of ectopic SAMs on the abaxial side of the leaves (McConnell and Barton, 1998; McConnell *et al.*, 2001). The class III HD-zip transcription factor genes *PHB, PHV* and *REV* are expressed on both the adaxial side of the cotyledons and in the embryonic SAM suggesting a role in the regulation of both SAM formation and the specification of adaxial cell fate (McConnell *et al.*, 2001).

The converse situation is observed in the overexpression of the abaxially expressed *FILAMENTOUS FLOWER (FIL)* and other *YABBY (YAB)* family genes, which induce an abaxialised leaf phenotype and cause the arrest of SAM development (Sawa *et al.*, 1999; Siegfried *et al.*, 1999). Therefore *YAB* family genes function to suppress adaxial cell fate and meristem formation, a function highlighted by the ectopic meristem production in *fil* and *yab3* mutants (Kumaran *et al.*, 2002).

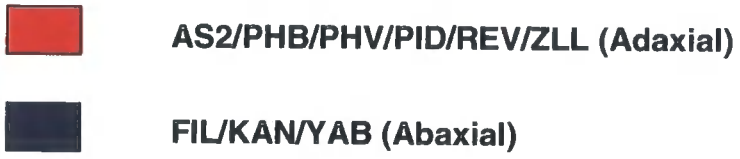
The abaxially expressed genes *KANADI1 (KAN1)* and *KANADI2 (KAN2)* encode members of the *GARP* family of transcription factors. In *kan1 kan2* double mutants, the majority of organs display an adaxial cell fate, and the expression domain of *PHB* is expanded, thus suggesting that *KAN* functions to promote abaxialisation and also to restrict *PHB* expression (Eshed *et al.*, 2001; Kerstetter *et al.*, 2001).

Taken together these observations suggest that the polarity of the cotyledons is intrinsically linked with the development of the SAM, with adaxial character

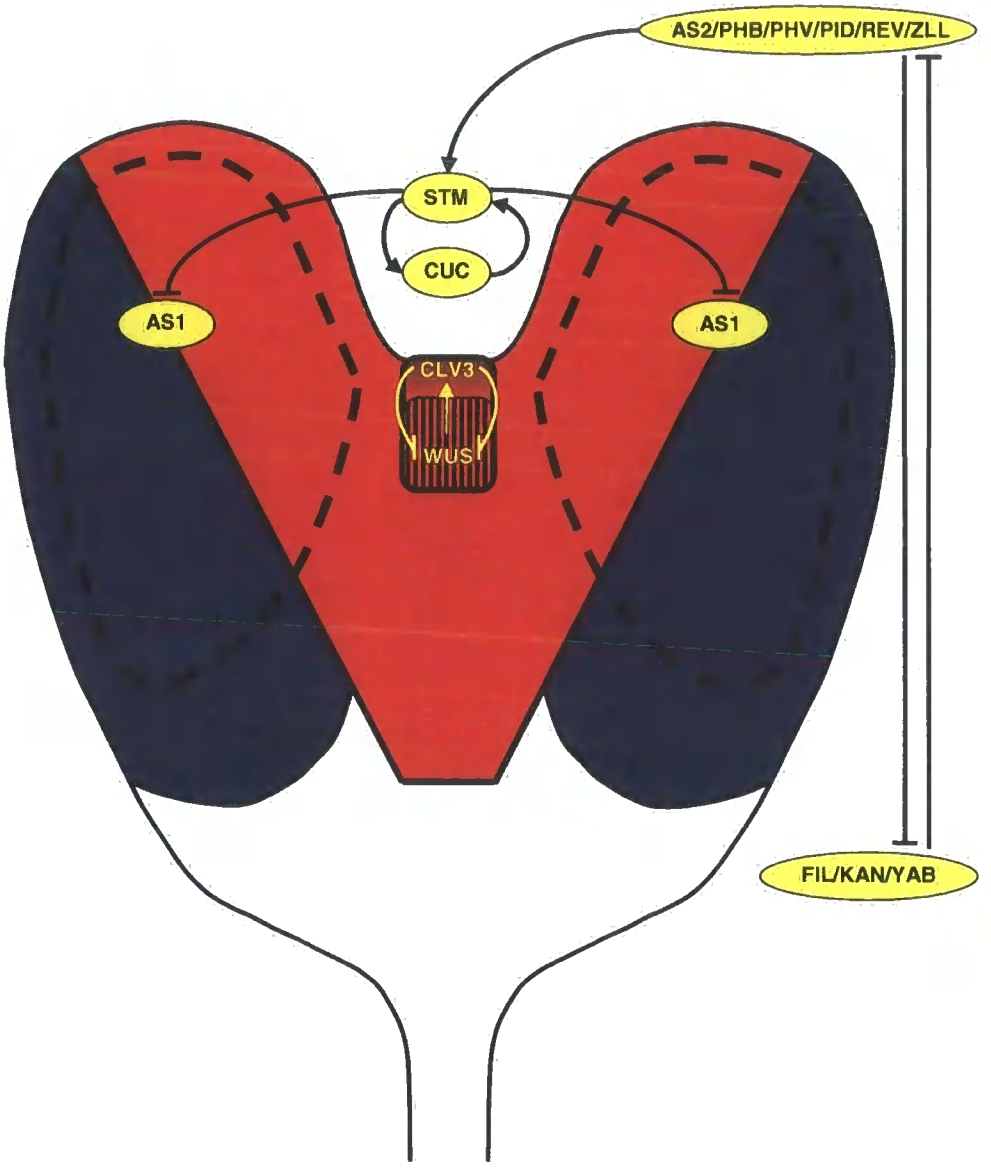
Figure 1.3 Genetic interactions among genes involved in SAM formation and abaxial/adaxial patterning of the cotyledons.

The size of the apical stem cell region is regulated via a feedback loop between *WUS* (positive) and *CLAVATA* genes (negative).

CUC genes positively regulate *STM* expression.
STM negatively regulates *AS1* expression in the presumptive SAM region.
AS1 is expressed throughout the cotyledon primordia and negatively regulates class I *KNOX* gene expression.



AS2, *PHB*, *PHV*, *PID*, *REV* and *ZLL/PNH* are expressed in the adaxial side of the cotyledon. The adaxial side is believed to promote correct SAM development.
KAN, *FIL/YAB1* and other *YAB* genes are expressed in the abaxial side of the cotyledon.
The adaxial and abaxial fate genes mutually repress each other.



Adapted from Takada and Tasaka, 2002, and Torres Ruiz, 2004.

promoting SAM formation while abaxial character represses it (see Figure 1.3). Reciprocal signalling is indicated by the classical experiments of Sussex (1954), whereby adaxial/abaxial polarity is lost when the influence of the SAM is removed (Sussex, 1954). This does not appear to be demonstrated in the *stm* mutant where the only cotyledonary abnormality is a small region of fusion at their base (Long *et al.*, 1996).

1.3.3.2 The central and basal regions

The hypocotyl and embryonic central region are derived from the lower portion of the lower tier of the embryo proper, a region that also contributes to root and root meristem. Mutational analysis highlighted the *FACKEL (FK)* gene in specification of the central region. The *fackel* mutant phenotype is apparent from the globular stage of embryogenesis and displays a specifically reduced hypocotyl, resulting from a failure to undergo asymmetric division. The resultant seedling phenotype appears to have cotyledons fused directly to a short root, although the presence of a greatly reduced hypocotyl has been demonstrated (Mayer *et al.*, 1991; Souter *et al.*, 2002). The *fackel* mutant has been independently characterised by three groups as being defective at an early stage in the sterol biosynthetic pathway (Jang *et al.*, 2000; Schrick *et al.*, 2000; Souter *et al.*, 2002). The *FACKEL* gene encodes a sterol C-14 reductase, with the mutant shown to have reduced levels of brassinosteroid (BR) hormones and the bulk sterols campesterol and sitosterol (Jang *et al.*, 2000; Schrick *et al.*, 2000).

Unlike the apical region the basal embryonic region is derived from two clonally distinct cell lines, the lower tier cells of the embryo proper originate from the apical zygotic cell, and the hypophysis originates from the basal zygotic cell. Signalling from the pro-embryo is believed to initiate the integration of these two cell lines (Jürgens, 2001).

The root meristem is a stem-cell system at the basal end of the seedling axis. It consists of a core of four mitotically inactive cells, known as the quiescent

centre, surrounded by two tiers of stem cell initials. The root meristem becomes active at the heart stage of embryogenesis, and the embryonic root begins to extend. The stem cell initials extend the embryonic root by giving rise to embryonic root tissue above and the central root cap below (Dolan *et al.*, 1993). A clonal boundary runs across the root meristem, with the root tissue initials derived from the apical zygotic cell exclusively compared to the quiescent centre and root cap initials which are established by the hypophysis (Dolan *et al.*, 1993; Scheres *et al.*, 1994). The failure to form a root meristem in *hobbit* (*hbt),* and other 'hypophyseal cell' mutants highlights the importance of clonal origin in the establishment of the root meristem. The initial defect in these mutants is aberrant development of the hypophysis, implicating its correct development as essential for subsequent root meristem development (Scheres *et al.,* 1996). The *hbt* mutation principally affects the hypophyseal precursor and results in a failure to form the quiescent centre (Willemsen *et al.,* 1998). Observation of such mutant phenotypes coupled with laser ablation studies suggest that the quiescent centre plays a role in the recruitment of adjacent cells as stem cells during embryonic root meristem formation, as well as maintenance of stem cell fate in the seedling (Van den Berg *et al.,* 1997; Willemsen *et al.,* 1998; Jürgens, 2001).

Mutations in the *MONOPTEROS (MP), BODENLOS (BDL)* and *AUXIN RESISTANT 6 (AXR6)* genes affect the embryonic root, but have been shown to be capable of forming roots post-embryonically (Berleth and Jürgens, 1993; Hamann *et al.*, 1999; Hobbie *et al.*, 2000). These genes therefore appear to be required for embryonic root formation, but not for root formation in general. Impaired signalling between the pro-embryo and the hypophysis appears to be implicated in the resultant mutant phenotype. Conversely to the *hobbit* phenotype, the proembryo in *mp* and *bdl* is initially abnormal before a defect is observed in the hypophysis (Berleth and Jürgens, 1993; Hamann *et al.*, 1999).

The *MP* gene is required for the formation of the hypocotyl, root, root meristem and the root cap (Berleth and Jürgens, 1993). Defective vascularisation of the cotyledons in the *mp* mutant also implicates it in the development of aligned

vascular strands and cell axialisation (Przemeck *et al.*, 1996). The mutant phenotype in *mp* embryos is observed from the eight-cell stage, at which point the *mp* embryo consists of four rather than two tiers of cells (Berleth and Jürgens, 1993). Subsequent aberrant division of the central cells results in a failure of the hypocotyl to elongate normally. The presumptive hypophysis undergoes suspensor-like division generating an aberrant clump of cells, rather than the basal region (Jürgens, 2001). This is suggestive that the hypophysis is extra-embryonic in nature and adopts an embryonic fate only in response to signals from the octant stage embryo (Mayer and Jürgens, 1998). The *MP* gene encodes an auxin response factor (ARF5), with the same binding specificity as *AUXIN RESPONSE FACTOR 1 (ARF1)* (Ulmasov *et al.*, 1997), a transcription factor that binds to auxin response elements (AREs) within the promoters of auxin-inducible genes (Guilfoyle *et al.*, 1998; Hardtke and Berleth, 1998). *MP* displays a similar expression pattern to *PIN1*, although it has been demonstrated that *PIN1* expression does not require *MP* gene function (Palme and Gälweiler, 1999; Steinmann *et al.,* 1999). Cell axialisation and vascular development is most likely mediated by *MP* through the transcriptional regulation of auxin-responsive genes (Souter and Lindsey, 2000).

The *bdl* mutant phenotype mirrors *mp* in that the apical cell divides horizontally rather than vertically and results in a double-octant proembryo (Berleth and Jürgens, 1993; Hamann *et al.*, 1999). Subsequent hypophyseal development is compromised, thus resulting in the lack of an embryonic root (quiescent centre and root cap). This mutant phenotype is also observed in *bdl mp* double mutants (Hamann *et al.*, 1999). *BODENLOS* encodes IAA12 (indole-3-acetic acid12), a short-lived nuclear protein, comprised of four conserved domains, the first IAA shown to be involved in embryogenesis (Hamann *et al.*, 2002). Domains III and IV are required for interaction with ARF transcription factors (Kim *et al.*, 1997; Reed, 2001; Rogg and Bartel, 2001). Transient assays in carrot protoplast demonstrated that IAA proteins (including IAA12) repress auxin-responsive gene activation (Tiwari *et al.*, 2001). Auxin relieves this repression through the promotion of binding between the IAA proteins and SCF^{TIR1} (a class of ubiquitin protein ligase), resulting in their ubiquitination and degradation (Gray *et al.*, 2001; Tiwari *et al.,* 2001; Zenser *et al.,* 2001;

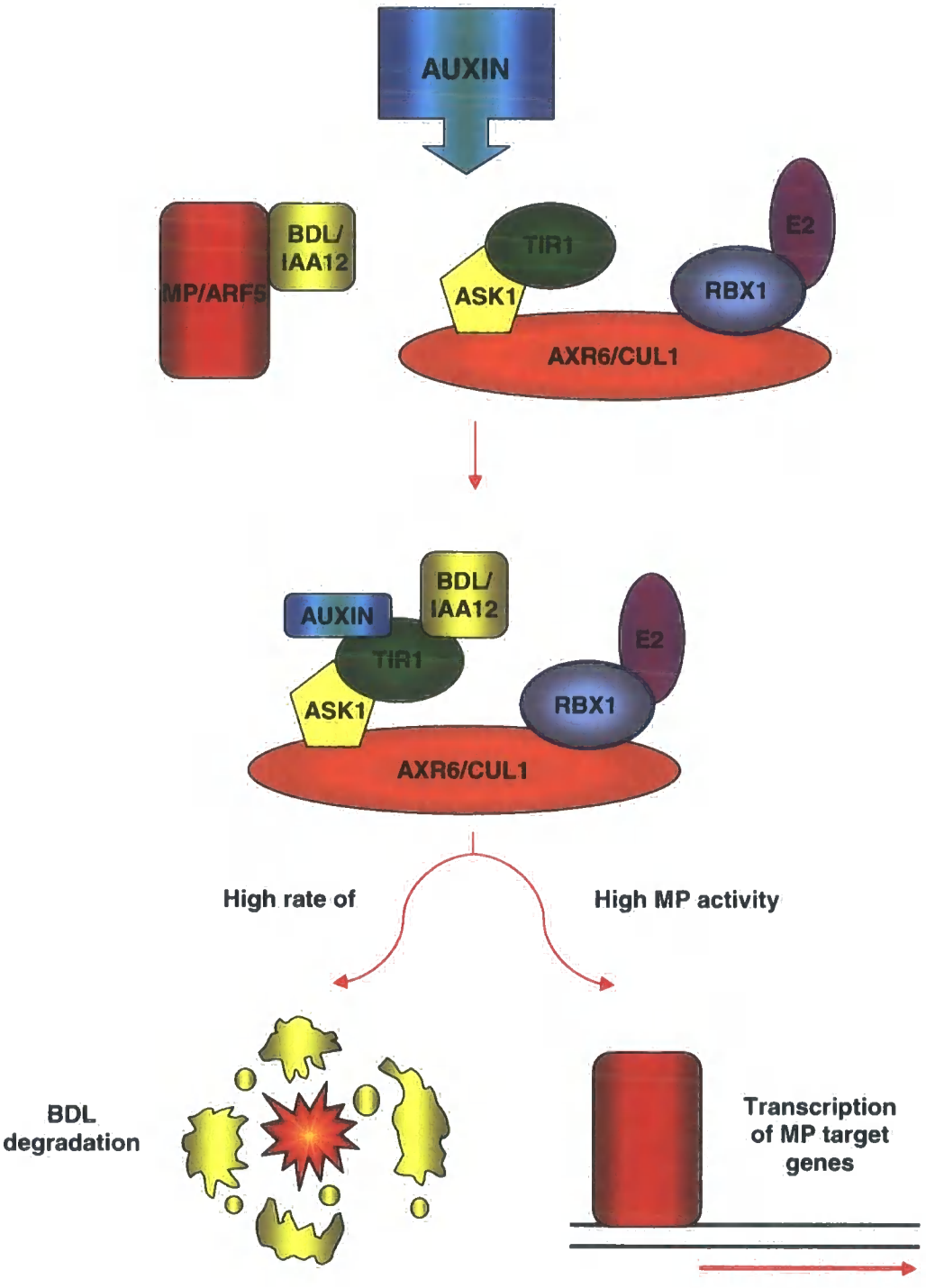
Dharmasiri *et al.*, 2005; Kepinski and Leyser, 2005). The mutant phenotype results from an alteration in the conserved domain II, which is involved in the ubiquitin-proteasome protein degradation pathway (Worley *et al.*, 2000; Gray *et al.*, 2001; Ramos *et al.*, 2001; Tiwari *et al.*, 2001). These observations suggest that a complex may be formed between *BDL* and *MP*, which prevents *MP* from activating target genes. In response to auxin the *BDL* protein is degraded, thus releasing *MP* (see Figure 1.4). Both *BDL* and *MP* are expressed in the proembryo but not the hypophysis. This suggests that an auxin-dependent signal must be relayed from the proembryo to the adjacent extra-embryonic cell in order to switch it to a hypophyseal cell fate (Hamann *et al.*, 1999).

auxin resistant6 (AXR6) mutants share phenotypic similarities with those of *mp* and *bdl*, but are novel in that cell division defects are found within both the embryo and the suspensor. The suspensor undergoes aberrant division to create radial layers rather than a single file. Consequently the hypophyseal cell does not form correctly, and the distinction between embryo proper and suspensor is lost. Development arrests shortly after germination, mutant seedlings lack both a root and a hypocotyl (Hobbie *et al.*, 2000). *AXR6* encodes the SCF subunit, CUL1, suggesting that mutation results in a failure to form a stable SCF^{TIR1} complex, and consequently reduced degradation of its substrate AXR2/IAA7. It has been proposed that the *axr6* mutation affects the levels of functional SCF resulting in an abnormal accumulation of *BDL* and thus repression of *MP* function (Hellmann *et al.*, 2003).

The *HOBBIT* gene encodes a subunit of the anaphase-promoting complex (APC), a class of ubiquitin protein ligase (Bilou *et al.*, 2002). Observation of the stabilisation of IAA17 in *hbt* mutants has led to the suggestion that IAA proteins are substrates of APC in the embryo. Therefore it seems possible that both APC and SCF^{TIR1} could play some role in the degradation of *BDL*/IAA12 during embryogenesis.

Figure 1.4 Auxin mediated interactions between *BDL*, *MP* and *AXR6*.

The plant hormone auxin acts to promote the degradation of the Aux/IAA transcriptional repressor proteins through the action of the ubiquitin protein ligase SCF^{TIR1}. Auxin binds directly to SCF^{TIR1} to promote its interaction with *BDL/IAA12*. Degradation of *BDL/IAA12* leads to the transcription of *MP* target genes. *AXR6* is an essential component of the functional SCF complex, when mutated it results in a stabilisation of the *BDL-MP* heterodimer and consequently reduced levels of *MP* activity.



Adapted from Hellmann *et al.*, 2003; Dharmasiri *et al.*, 2005 and Kepinski and Leyser, 2005.

1.4 Novel Methodologies: problems of high resolution transcriptional profiling at early stages and in specific domains

So far we have considered the approaches that have been employed with considerable success to isolate genes of importance in the development of the embryo. However, in the study of the regulation of gene expression such techniques are limited to the amount of data that can be obtained from a single experiment. In recent years the quantity of sequence data available for a range of organisms has increased rapidly. Combined with novel methodologies such as cDNA and oligonucleotide microarray this wealth of data has permitted the expression level analysis of a large number of genes in a single experiment, and given rise to the study of functional genomics (Brown and Botstein, 1999; Lockhart and Winzeler, 2000).

The analysis of global gene expression has the potential to reveal a great deal in terms of gene function and the interaction between genes throughout development. The cellular environment in which these developmental processes occur is a complex one with specific processes being carried out at a tissue or even cellular level. However, technical limitations have forced the majority of analyses to use whole organs, such as the leaf or root. Organs are composite structures comprised of complex tissues consisting of multiple cell types. As each cell type has a unique transcriptome, such analyses are in effect producing an averaged gene expression level across an organ, potentially masking genes of interest. Thus, in order to gain an understanding of the gene expression profile of an individual cell type, an efficient method is required to isolate populations of specific cell types. A number of such strategies have been employed.

1.4.1 Microcapillaries

Aspiration with microcapillaries has been shown capable of isolating specific cells from a living tissue. It has been employed to specifically sample tomato leaf epidermal or guard cells, from which mRNA was isolated for the

preparation of a cDNA library (Karrer *et al.*, 1995). Generally this methodology is limited to surface cells, although the expression of a green fluorescent protein (GFP) marker under the control of a tissue-specific promoter has allowed the identification of non-surface cells in transgenic plant lines (Sheen *et al.*, 1995; Brandt *et al.*, 1999, 2002). As relatively few such cell-specific promoters exist this method is limited.

1.4.2 Protoplast preparation

A second way to collect specific cell types is to prepare protoplasts from intact plant tissue. Protoplasts derived from cells of interest can be sorted using flow cytometry according to size or the presence of a cell-specific marker such as GFP (Galbraith *et al.*, 1988; Sheen *et al.*, 1995). Concerns over gene expression changes induced as a consequence of the manipulations involved in protoplasting, have raised doubts as to the application of this methodology in the gene expression profiling of specialised cell types (Grosset *et al.*, 1990). Birnbaum *et al.* (2003) used this method to generate a global expression map for the *Arabidopsis* root. They demonstrated that the protoplasting procedure did not dramatically alter the expression profile, but did isolate several hundred transcripts that were induced by the manipulations (Birnbaum *et al.*, 2003).

1.4.3 Laser capture microdissection (LCM)

Laser capture microdissection is a powerful tool allowing the rapid and precise isolation of specific populations of cells or even individual cells from a heterogeneous tissue, based on established histological identification (Emmert-Buck *et al.*, 1996; Bonner *et al.*, 1997; Simone *et al.*, 1998). Quick fixation or freezing of tissue samples for LCM minimises any undesirable changes in gene expression that could occur during sample preparation (Gillespie *et al.*, 2002). LCM has been utilised extensively in pathology and cancer biology, providing insight through the isolation of cancer cells from complex tumour tissue (Gillespie *et al.*, 2001) and is becoming of increasing importance in areas such as the study of neurodegenerative disease (Standaert, 2005).

To conduct LCM, a tissue section containing the cell type of interest is placed on a microscope stage. This is then covered with an isolation cap, which harbours a transparent thermoplastic polymer transfer film. An infrared (IR) laser beam is then targeted on the cells of interest, this activates the transfer film causing it to expand and impregnate the target cells, cementing the region of interest onto the cap. Due to the formation of a stronger adhesive force between the transfer cap and the tissue section than exists between the tissue section and the slide, the fused cells are torn out of the tissue section when the cap is lifted. This process can subsequently be repeated on new tissue sections to increase the number of cells acquired (see Figure 1.5).

Until recently animal tissue has been the exclusive subject for LCM, and protocols for fixation, sectioning and the extraction of RNA, DNA and proteins have been optimised for animal cells (Goldsworthy *et al.*, 1999). LCM is now beginning to be used in plant research, although structural and compositional differences between animal and plant cells have required the methodology to be adapted to account for the cell wall and vacuole. Protocols have now been developed in several plant species, including rice, maize and *Arabidopsis* (Matsunaga *et al.*, 1999; Asano *et al.*, 2002; Kerk *et al.*, 2003; Nakazono *et al.*, 2003; Liu *et al.*, 2004; Casson *et al.*, 2005; Inada and Wildermuth, 2005).

Of the recent studies in which LCM has been successfully implemented on plant tissue a range of technical problems have been overcome and the resultant plant total RNA used for a variety of purposes.

Asano *et al.* (2002) applied a modified LCM system to the isolation of phloem cells from rice leaf tissue. Total RNA was extracted from approximately 150 microdissected cells, and used to construct a cDNA library following T7 RNA polymerase amplification. A phloem specific putative amino acid permease was revealed by the study (Asano *et al.*, 2002). cDNA library construction from a microdissected papillary serous ovarian carcinoma lesion to identify anonymous ESTs has previously been applied in mammalian cells (Peterson *et al.*, 1998).

Nakazono *et al.* (2003) captured in excess of 10,000 epidermal, vascular bundle and bundle sheath cells from ethanol:acetic acid-fixed coleoptiles of maize. Extracted RNA was amplified with T7 RNA polymerase and hybridised to a microarray containing approximately 8800 maize cDNAs, approximately 3% of these were found to be preferentially expressed in epidermal cells or the vascular tissues. This study demonstrated the feasibility of combining LCM techniques with microarray analysis to conduct global gene expression analysis on plants.

Kerk *et al.* (2003) conducted LCM on paraffin-embedded tissue sections to preserve histological integrity compared to frozen tissue sections. RT-PCR was then used to examine gene expression patterns.

Casson *et al.* (2005), and this thesis, demonstrate the first application of LCM to fresh frozen plant embryonic tissue. Extracted RNA was amplified and hybridised to an Affymetrix ATH1 GeneChip® array. Approximately 65% of the genome was found to be expressed in the developing embryo at globular and heart stages (Casson *et al.*, 2005).

Inada and Wildermuth (2005) utilised a rapid microwave paraffin preparation method, requiring no fixative, for the preparation of *Arabidopsis* leaf tissue for laser microdissection. RT-PCR was then conducted on the extracted RNA to assess the specificity of harvested leaf cell types.

LCM is a powerful tool and one that lends itself particularly to the isolation of embryonic regions, which previously could not be accessed by conventional dissection techniques. A key to its application in transcriptional profiling is in the successful amplification of the small amounts of RNA retrieved from captured tissues.

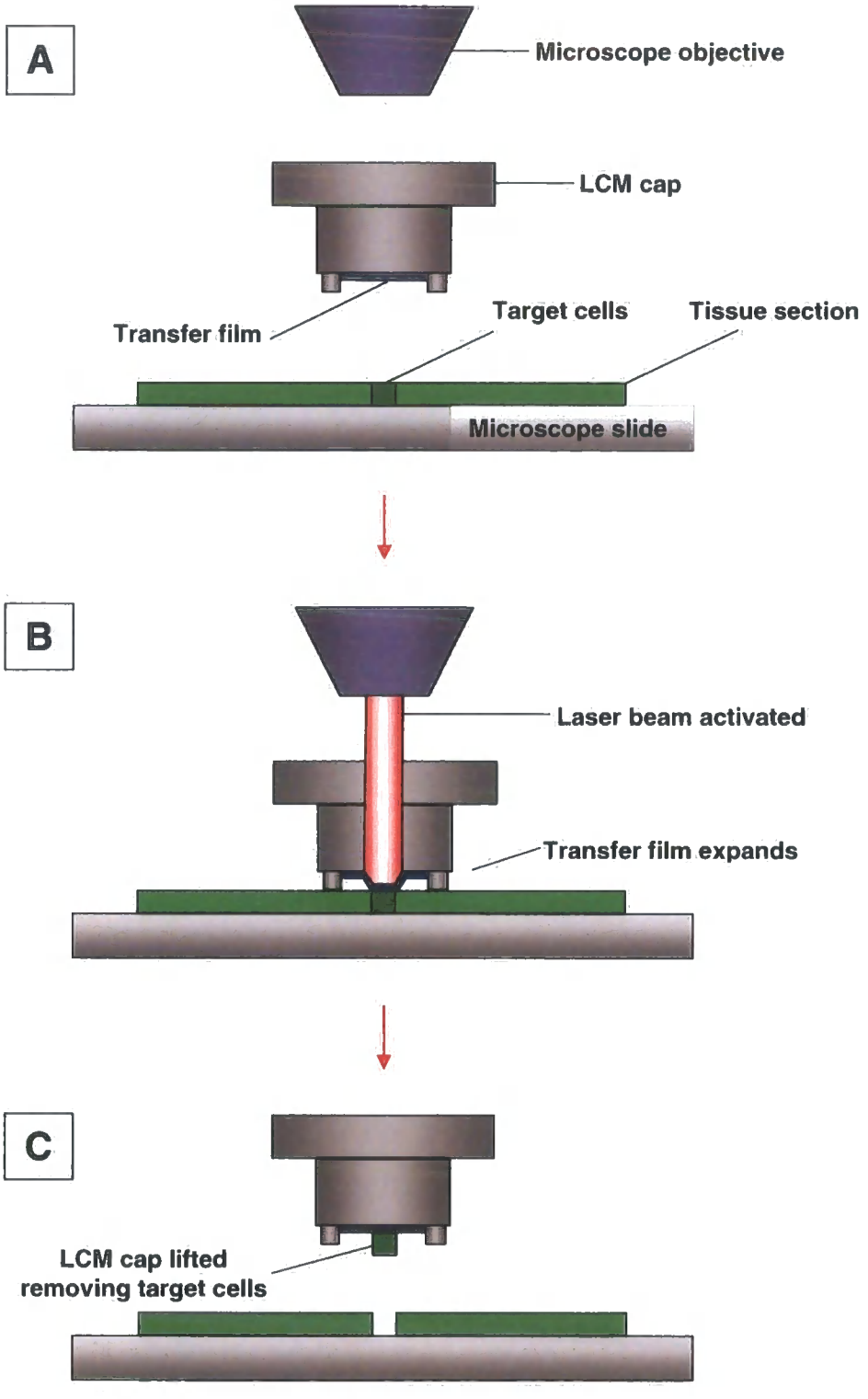
Figure 1.5 Laser Capture Microdissection (LCM).

A. A tissue section containing the cell type of interest is placed on a microscope stage. This is then covered with an isolation cap, which harbours a transparent thermoplastic polymer transfer film.

B. A focused laser beam is then targeted on the cells of interest, this activates the transfer film causing it to expand and impregnate the target cells, cementing the region of interest onto the cap.

C. Due to the formation of a stronger adhesive force between the transfer cap and the tissue section than exists between the tissue section and the slide, the fused cells are torn out of the tissue section when the cap is lifted.

This schematic describes the principles used by the PixCell II™ LCM system (Arcturus, CA, USA).


A
Microscope objective
LCM cap
Transfer film
Target cells
Tissue section
Microscope slide
B
Laser beam activated
Transfer film expands
C
LCM cap lifted
removing target cells

1.4.3.1 Linear RNA amplification

The introduction of microarray technology has ushered in a new level of transcriptome analysis. It provides an opportunity to study changes in gene expression over a time-course, between different tissues and between a mutant and its wild type background on a global level. A disadvantage to this technology when applied to the study of plant embryogenesis, particularly in combination with LCM, is the requirement for microgram quantities of RNA. Obtaining the amount of laser-captured embryonic tissue required for the isolation of µg quantities of RNA is unfeasible. Therefore amplification is required. However, in this case, this cannot be accomplished using conventional PCR, as there would be a resultant bias towards the smaller and more abundant transcripts in an mRNA population. For an analysis of relative gene expression levels between different populations this would obviously invalidate the results. To counteract this, a linear RNA amplification technique has been developed (using T7 RNA polymerase) by Van Gelder *et al.* (1990). Using this technique the relative abundance of individual mRNA sequences within an mRNA population is not distorted (Poirier and Erlander, 1998; Baugh *et al.*, 2001).

The linear RNA amplification procedure can be broken down into a number of steps as shown in Figure 1.6. Firstly, reverse transcription of the sample RNA is primed with an Oligo (dT) primer bearing a T7 RNA polymerase promoter to synthesise first strand cDNA. The synthesis of second strand cDNA involves the simultaneous action of RNase H to degrade the RNA and DNA polymerase to synthesise the second strand. This is followed by in-vitro transcription with a T7 RNA polymerase using the resultant double-stranded cDNA as a template. A single round of amplification can produce a 100-to-1000-fold increase in RNA yield. Further rounds can be performed until sufficient amplified RNA (aRNA) is available for analysis. Such protocols have successfully been used in combination with microarray technology to obtain the expression profile of single mammalian cells (Brady, 2000).

Figure 1.6 Linear RNA amplification overview.

The linear RNA amplification protocol described was developed in the laboratory of Dr. James Eberwine (Van Gelder *et al.*, 1990) and forms the foundation of kits such as the MessageAmp™ aRNA kit from Ambion (Europe Ltd; Huntingdon, UK).

First strand cDNA synthesis

Utilising the poly(A) tail of the mRNA by priming with a T7 Oligo (dT) primer, cDNA is synthesised harbouring a T7 promoter sequence.

Second strand cDNA synthesis

T7 cDNA generated by the previous step is converted into double-stranded DNA (dsDNA) to provide a template for *in vitro* transcription.

cDNA purification

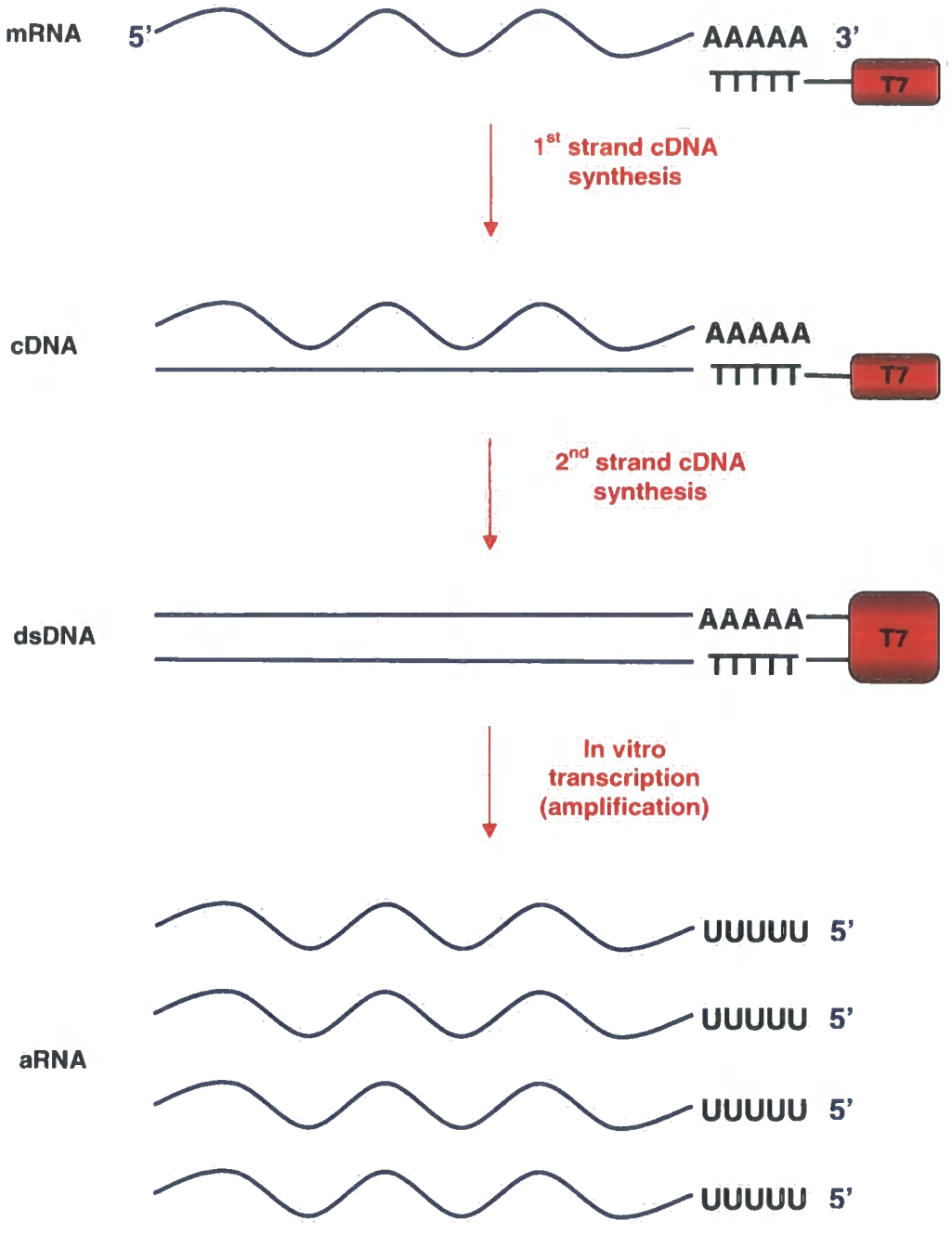
Removal of contaminants such as RNA, primers, enzymes and salts, which inhibit *in vitro* transcription.

***In vitro* transcription**

In vitro transcription of the purified dsDNA with T7 RNA Polymerase generates hundreds to thousands of antisense RNA (aRNA) copies of each mRNA in the sample.

aRNA purification

Removal of contaminants such as unincorporated NTP's, salts, enzymes and inorganic phosphate, which affect aRNA stability and inhibit subsequent enzymatic manipulation.



Adapted from MessageAmp™ aRNA kit instruction manual (Ambion, Austin, Texas, USA).

1.5 Microarray analysis of gene expression

Examination of gene expression levels can reveal distinct temporal changes occurring in induced cells or molecular differences between different tissues or even individual cell types. Microarrays are an incredibly powerful technique in that they allow the expression levels of thousands of genes to be simultaneously monitored in a single experiment.

The use of labelled nucleic acids to monitor expression level of nucleic acid molecules on a solid support was first reported thirty years ago (Southern, 1975). During the 1980's and early 1990's, several groups hybridised mRNA to cDNA libraries spotted onto nylon supports, with spacing down to approximately 2 mm (Lennon and Lehrach, 1991). The advent of the modern microarray came in the mid 1990's with the development of an automated spotting robot by the Pat Brown lab in Stanford University. This provided very accurate, reproducible spotted arrays, and engaged the interest of the wider research community (Schena *et al.*, 1995).

The general scheme which microarrays follow, is that mRNA extracted from the cell or tissue of interest is used to generate a labelled 'target', which is then hybridised simultaneously to a large number of DNA sequences, or 'probes', immobilised on a solid support. Thus allowing the expression levels of thousands of transcript species to be simultaneously quantified.

1.5.1 Array Platforms

The two most commonly used microarray systems are grouped according to the arrayed material: complementary DNA (cDNA) and oligonucleotide microarrays (see Figure 1.7).

On spotted microarrays, polymerase chain reaction (PCR) products generated from a cDNA clone set or a custom cDNA library are 'printed' onto glass slides or nylon membranes (Schena *et al.*, 1995; Duggan *et al.*, 1999). Genes are generally represented by a single DNA fragment several hundred base pairs in length (Harrington *et al.*, 2000). Spot sizes are generally between 100 and 300

Figure 1.7 Comparison of array preparation and expression assay between spotted DNA microarrays and high-density oligonucleotide arrays.

(a) Array fabrication and gene representation

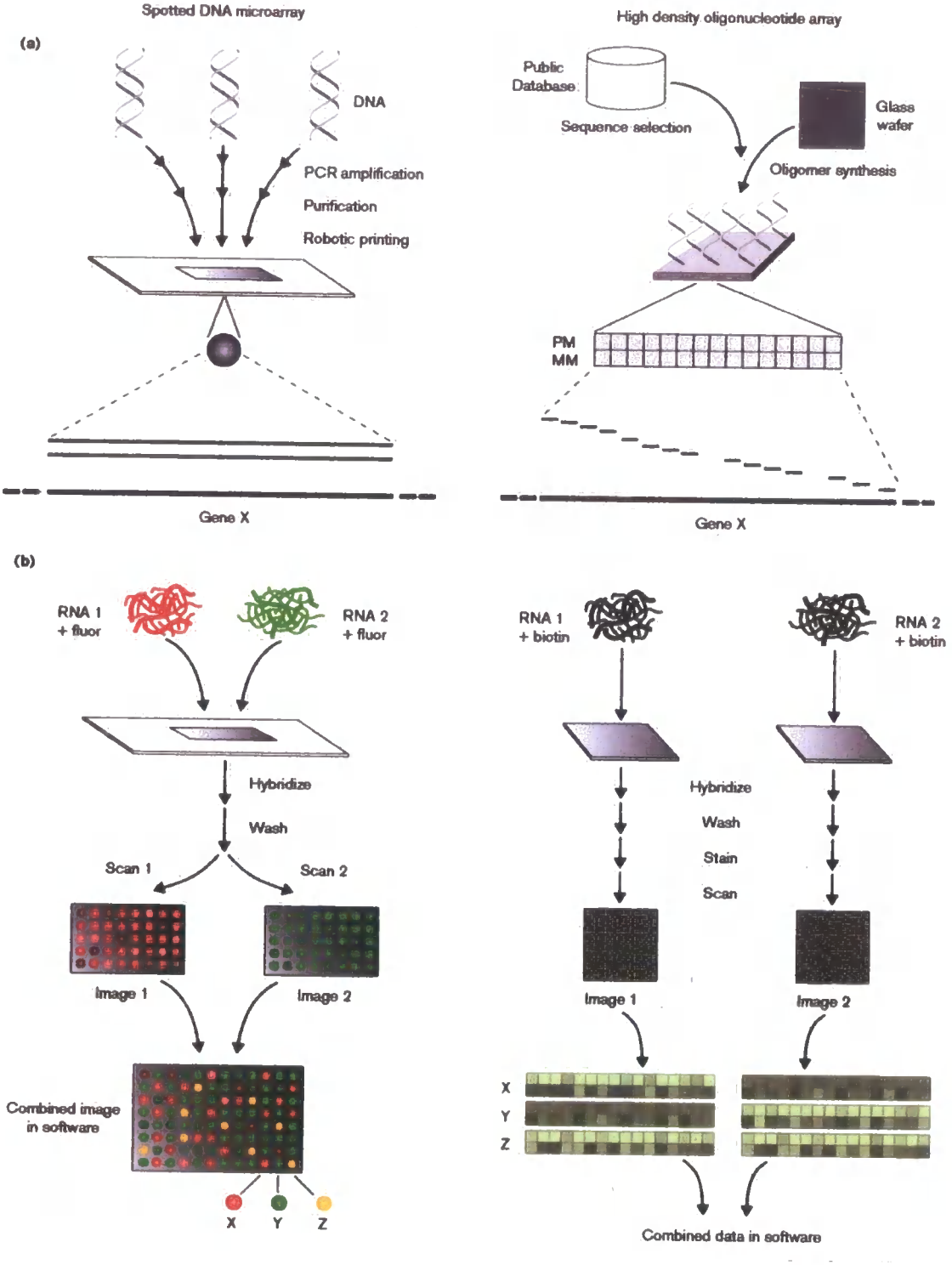
On spotted microarrays, polymerase chain reaction (PCR) products generated from a cDNA clone set or a custom cDNA library are 'printed' onto glass slides or nylon membranes (Schena *et al.,* 1995; Duggan *et al.*, 1999). Genes are generally represented by a single DNA fragment several hundred base pairs in length (Harrington *et al.*, 2000).

High density *in situ* synthesised oligonucleotide arrays are manufactured using a photolithographic manufacturing process developed by Foder *et al.* (1991), in which oligonucleotides are synthesised directly onto the glass surface. Every gene is represented by 11 different 25-mer oligonucleotide pairs (on the ATH1 GeneChip®) (PM, 'perfectly-matched' and MM, 'mismatched') spread across the chip.

(b) Expression assay

For spotted arrays, messenger RNA's from sample 1 (experimental) and sample 2 (control) are labelled with different fluorophores, hybridised to the probes on the surface of the array and subsequently scanned using each fluorescent dye's wavelength independently. Coloured dots labelled x, y and z at the bottom of the image correspond to hypothetical genes present at increased levels in sample 1 (x), sample 2 (y), and similar levels in both samples (z).

In the GeneChip® expression assay RNA is labelled to produce biotinylated cRNA. After hybridisation, bound biotin-cRNA is stained with a fluorophore conjugated to avidin, and detected by laser scanning. Sets of paired oligonucleotides for genes x, y and z are shown at the bottom.



Taken from Harrington *et al.*, 2000.

μm in diameter, with an approximately equivalent spacing between spots (Schultze and Downward, 2001). Arrays containing between 10,000 and 20,000 spots are commonly created on a standard glass microscope slide, although an array in the order of 80,000 spots is possible in this area (Burgess, 2001). Control and experimental cDNAs targets are differentially labelled with Cy3- and Cy5-dUTP fluorophores. The resultant labelled cDNA populations are then mixed and hybridised to a spotted array slide. Measuring the signal intensity for both fluorophores, and then calculating a ratio, can be used to determine the relative expression level of any transcript species between the two populations.

High density *in situ* synthesised oligonucleotide arrays are manufactured according to a method developed by Foder *et al.* (1991), they are now referred to as GeneChip® oligonucleotide arrays (Affymetrix Inc, Santa Clara, CA). This method is sequence driven and avoids the requirement for large clone libraries. It is particularly suited to the expression profiling of organisms with completed genome sequences. Affymetrix generate arrays using a photolithographic manufacturing process in which oligonucleotides are synthesised directly onto the glass surface (Lipshutz *et al.*, 1999). The array surface is coated with a photo-labile compound; this inactivates the nucleotide-binding sites, which are found at each position on the array. Exposure to light from a mercury lamp removes the photo-labile compound allowing the binding of nucleotides. To selectively target the light exposure a computer generated photolithographic mask is placed over the array, this has holes at appropriate positions. The nucleotides that bind to the de-protected positions are pre-linked to the photo-labile compound preventing additional nucleotides binding. Further masks can be applied to expose binding sites for the next nucleotide. Following in this stepwise fashion an oligonucleotide chain can be created. Individual elements of 5-10 μm in size can be created, although further miniaturisation is envisaged using such methods as non-linear semiconductor photo-resist technology (McGall *et al.*, 1996). The yield of the stepwise synthesis is approximately 95%, and as no purification is possible this limits the length of oligonucleotide chains produced to 25 bases (Stoughton, 2005). The potential of short oligonucleotides to result in less specific hybridisation and consequently reduced sensitivity, has led to the development of arraying using pre-synthesised 50 to 100-mer oligonucleotides (Kane *et al.*, 2000). The Affymetrix

GeneChip® arrays overcome potential reductions in specificity by representing every gene with between 11 and 20 different 25-mer oligonucleotide pairs spread across the chip (11 different 25-mer oligonucleotide pairs are utilised on the ATH1 GeneChip). An additional level of control on these arrays is the use of a 'perfect-match' (PM) and a 'mismatch' (MM) pair to determine the degree of non-specific binding. The MM oligonucleotide control is identical to its PM partner except for its 13th nucleotide (central), which is mutated and non-complementary to the transcript. This allows the level of cross-hybridisation and background signal to be estimated and subtracted from the PM.

The first *Arabidopsis thaliana* GeneChip® represented 8300 genes of the *Arabidopsis* genome (Zhu *et al.*, 2000). However, with the completion of the International Arabidopsis Sequencing Project (The *Arabidopsis* Genome Initiative, 2000), a single GeneChip® has become available representing 22,800 putative genes of the *Arabidopsis* genome, each represented by 11 probe pairs (http://www.affymetrix.com/products/arrays/specific/arab.affx). Thus, plant researchers working on *Arabidopsis* and related members of the *Brassicaceae* family have access to a very powerful tool in the elucidation of global patterns of gene expression changes between developmental stages or specific tissue types.

1.5.2 Target Preparation

In addition to the probe's used on the two major microarray systems, the preparation of the target is also fundamentally different, as a consequence of limitations in the array technology.

The general progression of the target through a microarray experiment follows a similar course in both systems. mRNA is extracted from cells or tissues of interest; it is then labelled whilst being converted to cDNA by reverse transcription. Labelled cDNA is then hybridised to the probes on the surface of the array, unbound cDNA is then washed off the array prior to the detection of a signal.

The process of gridding used on spotted arrays is currently not accurate enough to allow comparisons to be made between signals generated from independent arrays. Therefore each slide represents a unique comparison between a control and an experimental population. Control and experimental cDNAs targets are differentially labelled with fluorescent dyes (such as Cy3- and Cy5-dUTP). The two, labelled cDNA populations, are then mixed and hybridised to the same array, this results in competitive binding of the targets to the probe sequences. Following hybridisation and washing, the array is scanned using each fluorescent dye's wavelength independently. A ratio can be calculated by comparing the signal intensity of the same spot under both wavelengths. In this way, the relative expression level of any transcript species between the two populations can be determined (Harrington *et al.*, 2000; Burgess, 2001; Schulze and Downward, 2001).

The *in situ* synthesis used on oligonucleotide chips is highly reproducible and therefore allows accurate comparisons to be made between signals generated from independent arrays. The labelling process for the GeneChip® array requires mRNA to be converted into biotinylated cRNA via an oligo-dT-primed cDNA intermediate. Each sample is then hybridised to a separate array. After washing, Streptavidin-Phycoerythrin is used to stain the biotin-bound probes. Phycoerythrin is detected at a wavelength of 570 nm by phospho-imaging. Transcript expression levels are then calculated according to the concentration of known positive control cRNA, spiked in to the hybridisation mixture. The expression level of transcript species from different experimental samples hybridised to the same type of array can then be compared (Harrington *et al.*, 2000; Burgess, 2001; Schulze and Downward, 2001; Stoughton, 2005).

As a powerful but ultimately expensive technology, microarrays were used initially (and still are in some cases) without sufficient replication of measurements. The use of replicates has been demonstrated to significantly reduce the number of potential false positive results, and is required to establish a high degree of confidence in the data.

Variability in microarray results has been shown to be of significance, particularly in the case of low-abundance transcripts, which may be of

considerable importance yet whose expression is not much above the background level. Negative controls such as the Affymetrix PM-MM go some way towards determining the level of background. Two-colour competitive hybridisations intrinsically normalise and correct for noise and background in a pairwise comparison. However different pairings of spotted arrays can only be created *in silico* using ratios of ratios, which can lead to error (He *et al.*, 2003).

1.5.4 Microarray experimental approaches

The success of microarray experiments when applied to a biological problem is dependent on the researcher having a clear question(s) and tailoring their approach to this. The two extremes of experimental approach to analyse gene expression are local and global, generally experiments fall somewhere between.

The local approach concerns itself with identifying single gene changes that might be responsible for the phenotype under investigation. An example of such an approach may include a comparative study between two biological conditions, such as a disease state and a wild-type. Genes up-regulated or down-regulated, in the disease state can be seen as more likely than random chance to be involved in the disease process, and as such could play a role in elucidating the molecular phenotype of the disease. Such a study proved successful in identifying genes highly expressed in a cancerous tissue compared to healthy tissue, providing potential drug targets (Hansel *et al.*, 2003). However, whilst the technology is capable of uncovering many putative candidate genes that have an altered expression, it cannot mark any of these as being of particular significance in the generation of a phenotype. To validate the approach candidate genes identified must then be analysed to assess the degree to which they contribute to the resultant phenotype, if at all. Conventional techniques such as overexpression, gene knockout and RNA interference can all be used to achieve this aim. Potential obstacles to the analysis of individual pathways include the potential functional redundancy of related gene products and the degree to which a single gene in a complex pathway is sufficient to induce a phenotype. While all studies must assume that

the mRNA transcript level is a sufficient indicator of biological control, further levels of complexity such as the post-translational modification of protein activity could hold the key to the eventual phenotype.

A more complex study is represented by the global approach, where global as opposed to single gene expression patterns are analysed. Microarrays experiments lend themselves very well to the analysis of genes with similar expression patterns when combined with advanced software packages capable of clustering genes. Large-scale analyses into the co-regulation of genes to reveal functional gene groups (termed 'synexpression groups') have been successfully undertaken in higher eukaryotes (Niehrs and Pollet, 1999). Global analyses have followed gene expression changes through the diauxic shift of the cultured yeast *Saccharomyces cerevisiae*. Genes in related metabolic pathways showed similar expression over the time course, comparable transcriptional changes observed in other genes provided a wealth of extra detail to the pathways (DeRisi *et al.*, 1997). Gene expression changes through the yeast and mammalian cell cycles have identified many new genes regulated by the cell cycles to those already known, and allowed functional relationships to be suggested between different points of the cycles (Cho *et al.*, 2001; Spellman *et al.*, 1998). Studies of the progression of gene expression through development in both *Drosophila* and *Caenorhabditis elegans* have yielded strong co-regulated gene groupings (White *et al.*, 1999; Kim *et al.*, 2001). In addition to common transcriptional control, it can be inferred that genes sharing a similar expression pattern are functionally related through 'guilt by association' (Clare and King, 2002). Hughes *et al.* profiled hundreds of different single-gene disruption mutants in yeast, compared to a wild-type control, and combined the data sets into a compendium database. The resulting patterns of the major pathway groupings allowed functional inferences to be made with considerable accuracy for previously un-annotated genes with similar expression profiles (Hughes *et al.*, 2000). The analysis of huge standardised datasets will hopefully lead to a greater understanding of the operation of the cell.

Gene expression analyses using microarrays on plant tissue is a rapidly developing area. Microarrays have already been utilised in the characterisation

of genes important in the regulation of plant defence mechanisms, oxidative stress responses, nitrate assimilation, phytochrome signalling, fruit ripening and seed development (Aharoni and Vorst, 2002). The sequencing of the *Arabidopsis thaliana* genome largely set the stage for global gene expression studies in plants, now the availability of the *Arabidopsis* ATH1 GeneChip® to the plant community at a non-prohibitive price promises to see a continued expansion to the already considerable wealth of data.

1.6 Project Aims and Objectives

The aim of this work was to develop a protocol for the use of LCM on torpedo-stage embryos. Using this technology to capture tissue from the cotyledon, root and shoot apical meristem regions, RNA could be extracted from this tissue and amplified to allow its use in downstream applications.

Providing good quality RNA could be obtained from embryonic tissue the next step would be to apply microarray technology using the *Arabidopsis* ATH1 GeneChip® to allow a global gene expression analysis. Data obtained would be validated using a variety of approaches including literature searches, RT-PCR and analysis of promoter-GUS fusions for putative transcription factors, of potential importance in the formation of apical-basal pattern.

Following validation the data would be subjected to high-level computational analysis to identify genes, which exhibited significant differential expression between tissue types.

Thus, the overall aim of this work was to develop a protocol for the application of the novel technology of laser-capture microdissection to embryos of the model plant species *Arabidopsis thaliana*, thus allowing the problem of the genetic control of apical-basal patterning to be addressed through the subsequent application of microarray technology and bioinformatic analysis.

Chapter 2

Materials and Methods

2.0 Materials and Methods

This chapter describes the materials and methods that were used to obtain the results as described in subsequent results chapters.

2.1 Materials

2.1.1 Chemicals

The chemicals used for this research were obtained from; Sigma Chemical Company Ltd. (Poole, UK), Fisher Scientific (Loughborough, UK), Fisons Scientific Equipment (Loughborough, UK), VWR (Poole, UK), Duchefa Biochemie (Haarlem, The Netherlands), BD (Sparks, MD, USA), or Bio-Rad Laboratories (Hemel Hempstead, UK) unless otherwise stated.

X-Gluc was from Melford Laboratories (Suffolk, UK). X-Gal was from Bioline (London, UK).

2.1.2 Enzymes

T4 DNA ligase and Superscript™ III reverse transcriptase were obtained from Invitrogen (Paisley, UK). Restriction endonucleases were obtained from Fermentas (Burlington, Canada) or Promega (Southampton, UK). Shrimp alkaline phosphatase and RNasin ribonuclease inhibitor were obtained from Promega Ltd. (Southampton, UK). Taq DNA polymerase was from Bioline. The Expand™ High Fidelity PCR system was from Roche (Lewes, UK).

2.1.3 Oligodeoxynucleotides

Oligodeoxynucleotide primers used in PCR reactions were obtained from MWG-Biotech (Ebersberg, Germany). Random Primers were obtained from Promega (Southampton, UK).

2.1.4 Kits

The GenElute™ plasmid miniprep kit and the GenElute™ plant genomic DNA kit were from Sigma. The *High Pure* PCR Clean-Up kit was from Roche. The TOPO-TA cloning kit was from Invitrogen (Paisley, UK). The plasmid midi kit DNA purification system, the RNeasy Plant RNA extraction kit and the QIAquick gel extraction kit were from Qiagen Ltd. (Crawley, England). The MessageAmp™ II aRNA kit was from Ambion (Europe) Ltd (Huntingdon, UK). The Absolutely RNA™ Nanoprep kit was from Stratagene (CA, USA).

2.1.5 Bacterial strains

The *E.coli* strains used were as follows: The TOP10 strain was a host for pCR®2.1-TOPO (Grant, S.G.N. *et al.*, 1990). The DH5α strain was used to prepare competent cells and as a plasmid host.

Agrobacterium tumefaciens C58C1 (Koncz and Schell, 1986) was used for plant transformations. This strain has been disabled so that it does not cause crown-gall disease but it still has the virulence factors required for T-DNA insertion into plant genomic DNA. Selection for the C58C1 strain was with 25 μg/ml nalidixic acid.

2.1.6 Plasmids

The following plasmids were used during this project: pCR®2.1 TOPO (Figure 2.1; Invitrogen) and pΔGUS-CIRCE (Figure 2.2). pCR®2.1 TOPO is used for cloning DNA fragments generated by PCR. pΔGUS-CIRCE was used in the construction of promoter-GUS constructs and contains the *β-glucuronidase* gene followed by the NOS terminator. pΔGUS-CIRCE is a wide host range binary cloning vector for *Agrobacterium*-mediated gene transfer into plant cells.

Figure 2.1 pCR® 2.1 TOPO

f1 Origin

pMB1 Origin

LacZα fragment containing TOPO® Cloning site

Kanamycin resistance ORF

Ampicillin resistance ORF

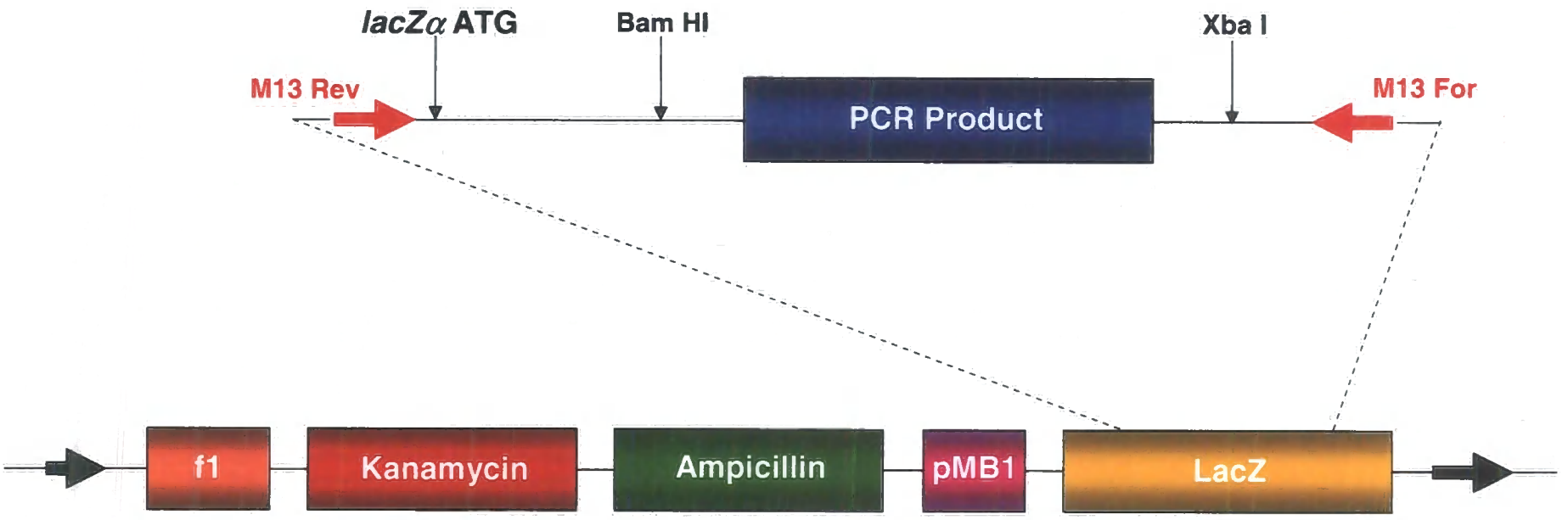
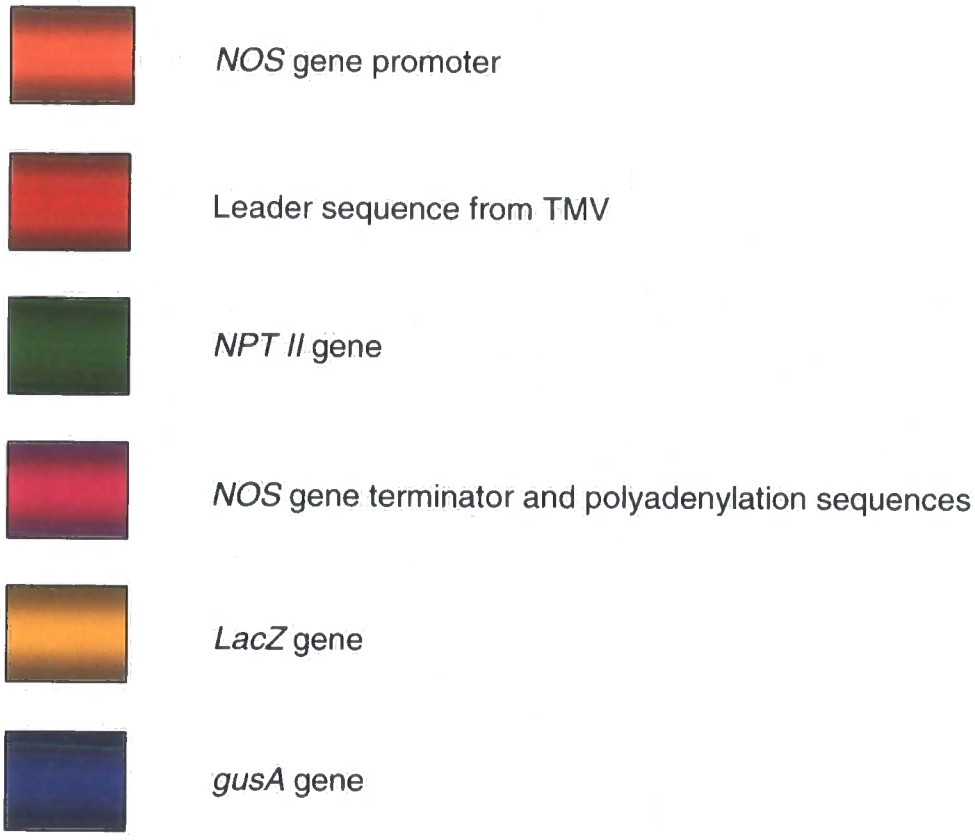
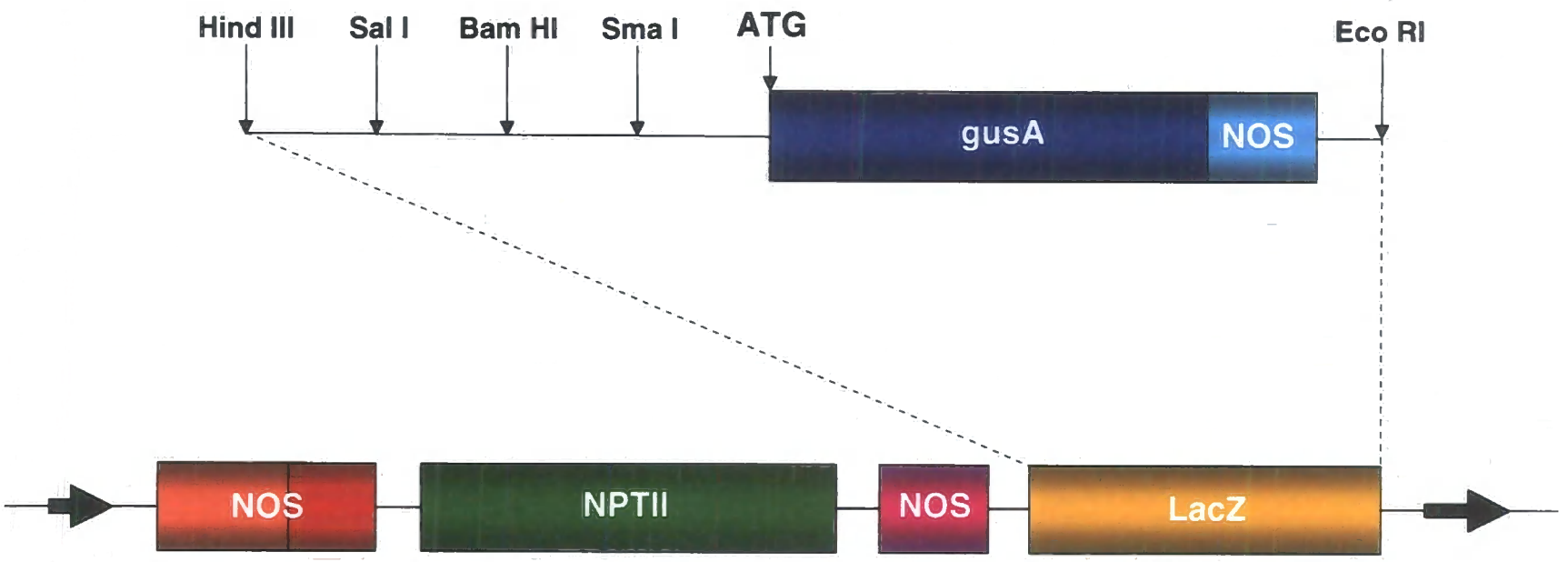

lacZα ATG
Bam HI
Xba I
M13 Rev
M13 For
PCR Product
f1
Kanamycin
Ampicillin
pMB1
LacZ

Figure 2.2 pΔGUS-CIRCE




Hind III
Sal I
Bam HI
Sma I
ATG
Eco RI
gusA
NOS
NOS
NPTII
NOS
LacZ

2.1.7 Bacterial culture media

All bacterial culture media was sterilized by autoclaving for 20 minutes at 121°C.

- **Luria-Bertani (LB) medium**: 10 g/l Bacto-tryptone, 5 g/l Bacto-yeast extract, and 5 g/l NaCl.
- **LB agar**: As for LB medium with 15 g/l Bacto-agar added prior to autoclaving.
- **SOC medium**: 20 g/l Bacto-tryptone, 5 g/l Bacto-yeast extract, and 5.84 g/l NaCl. pH adjusted to 6.8-7.0 and autoclaved. 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 2 g/l filter-sterilised glucose, added after autoclaving.

2.1.8 Plant material

Arabidopsis thaliana ecotype Col-0 was used for all experiments. T-DNA insertion lines were obtained from the Nottingham Arabidopsis Stock Centre as detailed in Table 2.1 (Nottingham, UK; NASC on-line catalogue, http://arabidopsis.info; Alonso *et al.*, 2003).

2.1.9 Plant culture media

- **1/2 MS10**: Half strength Murashige and Skoog medium (Murashige and Skoog, 1962), 10 g/l sucrose, adjusted to pH 5.8 with 1 M KOH, 8 g/l bactoagar added and autoclaved at 121°C for 20 minutes.

2.1.10 Antibiotics

Antibiotic stocks were filter-sterilised using 0.2 μm acrodiscs (Pall Corporation. Ann Arbor, MI, USA) and added to molten 1/2 MS10 or LB (cooled to approximately 60°C) and mixed thoroughly prior to pouring.

Table 2.1 ***Arabidopsis thaliana*** **lines**

Line Name	Source	Background	Description
Columbia (Col-0)	Lehle Seeds	Columbia	Wild Type
SALK_012862	NASC	Columbia	T-DNA Insertion
SALK_033446	NASC	Columbia	T-DNA Insertion
SALK_063571	NASC	Columbia	T-DNA Insertion
SALK_068359	NASC	Columbia	T-DNA Insertion
SALK_068825	NASC	Columbia	T-DNA Insertion
SALK_106430	NASC	Columbia	T-DNA Insertion
SALK_116644	NASC	Columbia	T-DNA Insertion
SALK_131239	NASC	Columbia	T-DNA Insertion
SALK_132004	NASC	Columbia	T-DNA Insertion
SALK_135981	NASC	Columbia	T-DNA Insertion
SALK_144520	NASC	Columbia	T-DNA Insertion
SALK_508845	NASC	Columbia	T-DNA Insertion

2.2 Bacterial Growth Conditions

E. coli was grown in LB media (with shaking) or on LB agar plates at 37°C.
A. tumefaciens was grown in LB media (with shaking) or on LB agar plates at 30°C.

For long-term storage bacterial strains were frozen at either -20°C or -80°C in the form of glycerol stocks, made by adding 0.5 ml of an overnight liquid culture to 0.5 ml of filter-sterilised 50% glycerol.

2.3 Plant Growth Conditions

2.3.1 Plants germinated on culture media under sterile conditions

Culture Media

For sterile culture of plant material, half-strength Murashige and Skoog medium was used; 2.2 g/l Murashige and Skoog basal mix, 10 g/l sucrose and 8 g/l of Bactoagar at pH 5.8 (1 M KOH). Agar was autoclaved at 121°C for 20 minutes before use, and then poured under sterile conditions in a sterile airflow cabinet.

Seed Sterilisation, Plating-out and Germination

It is necessary to sterilise seeds to remove any contaminating fungi, yeast or bacteria, which would otherwise thrive on the nutrient-rich culture medium.

Seeds were sterilised and washed in 70% ethanol (30 seconds), 20% hyperchlorite bleach/ 0.2% Tween 20 solution (30 minutes), and autoclaved distilled water (three washes). This sterilisation was conducted in a sterile airflow cabinet. Sterile seeds were then transferred to sterile Petri dishes (SLS. Nottingham, UK) containing half-strength MS media, with sterile disposable pipettes. Plates were labelled and then sealed with 3M Micropore™ surgical tape (3M Health Care, Germany). To increase germination frequency and

ensure relative uniformity of germination timing, plates were wrapped in aluminium foil before placing into the cold store (at 4°C) for 4 days to vernalise.

Sterile growth room conditions

The tissue culture growth room was at a temperature of 22 ±2°C, with a 16-hour light regime, at a photon flux density of 150 μmol m^{-2} s^{-1}.

Transfer of seedlings under sterile conditions for primary root length assays

Seeds were surface sterilised as described above and allowed to germinate on horizontal plates. Seedlings were then transferred to fresh plates three days after germination. Secondary plates were then incubated vertically (>60°), so as to allow the measurement of primary roots. Approximately 12 seedlings were transferred to each secondary plate to avoid overcrowding.

Transfers were undertaken in a sterile Flow hood, with sterile conditions maximised through the ethanol washing and subsequent flaming of forceps.

Primary root length assays

The main root exhibits a predictable level of growth and was therefore chosen to study in all root growth experiments. The length of the primary root was measured every three days. The same plants were used for all measurements in a developmental time-course.

2.3.2 Plants transferred to Soil

Soil was made up to a mixture of 6:1 Levington's multipurpose compost to silver sand (for enhanced drainage). Soil and sand were obtained from Klondyke Garden Centre, Chester-le-Street, UK.

7-14 days after germination on half strength MS media, young seedlings were transferred to individual pots containing soil.

Embryo sacs were obtained from siliques of 3-4 week old plants as described in Section 2.5.1

Aracon tubes (BetaTech. Gent, Belgium) were used to aid bulk seed collection.

Pest Control Measures

Soil was initially treated with a systemic insecticide called Intercept^{TM} (Scotts, Ipswich, UK) at a concentration of 0.2 g L^{-1} in tap water. Intercept^{TM} is an insecticide that contains 70% (w/w) imidacloprid that is used as a one-off soil drench. Intercept^{TM} was used to control sciarid fly, greenfly, whitefly and vine weevil. Biocontrol of thrips was also used in the growth rooms, through the use of mites (Syngenta).

Greenhouse Conditions

The green house had a light regime of 16 hours light (at a temperature of 22 ±3°C) to 8 hours dark (at a temperature of 15°C). Plants were watered daily from above, using a watering can with a fine nozzle.

2.4 GUS enzyme analysis

2.4.1 Histochemical localisation of GUS activity

(After Stomp, 1990).

Solutions

- **X-Gluc stock**: 20 mM X-Gluc (5-Bromo-4-Chloro-3-Indolyl-β-D-Glucuronide) in N-N-dimethylformamide, stored at -20°C.
- **X-Gluc buffer**: 100 mM NaH_2PO_4, 10 mM EDTA, 0.1% (v/v) Triton X-100, 1 mM potassium ferricyanide ($K_3[Fe(CN)_6]$) and 1 mM potassium ferrocyanide ($K_4[Fe(CN)_6]$), pH7.0.
- **X-Gluc staining solution**: prepared by mixing 1 volume of X-Gluc stock with 19 volumes of X-Gluc buffer to give a final concentration of 1mM X-Gluc.
- **Chloralhydrate solution**: 8 g chloralhydrate ($Cl_3CCH(OH)_2$), 3 ml water, and 1 ml glycerol.

Method

β-glucuronidase (GUS) enzyme activity, and thus by association *gusA* reporter gene activity, can be directly visualised within specific cells and tissues of transformed plants in a histochemical fashion using the chromogenic substrate 5-bromo-4-chloro-3-indoyl-β-D-glucuronic acid (X-Gluc; Melford Labs, Ipswich, UK. Jefferson, 1987). The effect of GUS activity on this substrate is to cause oxidative dimerisation to form an easily visualized, insoluble blue dye. The addition of the oxidative catalysts potassium ferricyanide and ferrocyanide acts to inhibit the diffusion of a soluble reaction intermediate into surrounding tissues (Mascarenhas and Hamilton, 1992).

Samples were vacuum infiltrated for 10 min in X-Gluc staining solution before incubation at 37°C for a variable time period, depending on the GUS activity level expected in the given sample.

2.4.2 GUS Assay

The optimal staining time was determined through a time-course experiment. Lines were stained for different lengths of time, 0.5, 1, 2, 3, 4, 5, 6, 7, 8 and 24 hours. At each time point the staining was analysed using light microscopy.

Once optimal staining was achieved, lines were cleared of chlorophyll to aid visualisation of the blue precipitate, by replacing the X-Gluc staining solution with 70% (v/v) ethanol for 24 hours, with regular changes of clean ethanol. Once cleared, seedlings were mounted in chloral hydrate for microscopic analysis as described in Section 2.5.

2.5 Microscopy and photography

Photographs were taken using a CoolSNAP *cf* digital camera (Photometrics, Roper Scientific, Tucson, AZ, USA) with Openlab 3.1.1 software (Improvision Ltd, Coventry, UK), on Leica MZ125 (Leica Microsystems (UK) Ltd, Milton Keynes, UK), Olympus SZH10 (Olympus UK Ltd, Southall) or Zeiss Axioskop (Carl Zeiss Ltd, Welwyn Garden City, UK) microscopes. Images were processed in Adobe Photoshop 5.0 (Mountain View, CA, USA).

2.6 Laser Capture Microdissection

2.6.1 Sectioning

Embryo sacs containing torpedo stage embryos were dissected from siliques and embedded in OCT embedding medium (10.24% polyvinyl alcohol, 4.26% polyethylene glycol, 85.50% Non-reactive ingredient; RA Lamb, Eastbourne, UK) in a base mould (Merck Eurolab, Dorset, UK). Samples were then immediately frozen in liquid nitrogen-cooled isopentane, prior to sectioning. 8 µm sections were cut on a Leica CM3050S cryostat (Leica Microsystems, Nussloch, Germany) at –22°C. Sections were collected on RNase-free uncoated glass slides (slides were prepared by cleaning with acetone followed

by baking at 200°C) and stored in 70% ethanol at -22°C ready for processing. Sections were processed through the following ethanol dehydration series: 30 sec in 70% ethanol, 15 sec in 0.01% phosphate buffered saline (PBS) pH 7.4, 30 sec in 70% ethanol, 30 sec in 95% ethanol, 30 sec in 100% ethanol, 2x10min in xylene. Slides were air dried and kept dry in boxes containing silica gel (Sigma).

2.6.2 Laser Capture Microdissection

LCM was performed using a PixCell II™ system using CapSure™ HS LCM caps (Arcturus, Ca, USA) as shown in Figure 1.5. Sample slides with dehydrated sections (see Section 2.6.1), were prepared for LCM using a prep-strip™ (Arcturus) to remove loosely adhered detritus. A slide was placed into position on the LCM microscope stage. A HS LCM cap (which harbours a transparent thermoplastic polymer transfer film) was lowered onto the section using the placement arm. The infrared (IR) laser beam was then set to a diameter of 7.5 μm and focused. Cells of interest were then located and the laser fired, this activates the transfer film causing it to expand and impregnate the target cells, cementing the region of interest onto the cap. The laser beam was set to a diameter of either 7.5μm for meristem captures or 15 μm for root and cotyledon captures. The power and duration of the laser beam was variable and dependent on beam size selected, but was typically 85 mW and 100 ms for a 7.5 μm beam. Captured cells were then removed from the parent tissue by rapidly lifting the LCM cap using the placement arm. Due to the formation of a stronger adhesive force between the transfer cap and the tissue section than exists between the tissue section and the slide, the fused cells are torn out of the tissue section when the cap is lifted. Typically, around 15 embryo sections (each sampled multiple times to achieve greater capture) were processed per LCM cap and non-specific material was removed from the surface of a LCM cap using a sterile Post-It note. It is estimated that approximately 200-300 cells were captured per LCM cap, this was not accurately quantified. An ExtracSure™ Sample Extraction Device (Arcturus) was then attached to the LCM cap in preparation for RNA extraction.

2.6.3 RNA Extraction

RNA from LCM cells was extracted using the Absolutely RNA™ Nanoprep kit (Stratagene, CA, USA) according to the manufacturers' instruction with slight modifications. In brief, 100 µl of lysis buffer and 0.7 µl β-mercaptoethanol was applied to the captured cells on the cap via the ExtracSure™ Sample Extraction Device, which was then connected to a 0.5 ml microcentrifuge tube. After vortexing the sample was incubated at 60°C for 5 min. The lysis buffer was then collected and mixed with an equal volume (~100 µl) of 70% ethanol before applying to an RNA binding column. Following DNase treatment (to remove contaminating gDNA) and a series of washes, RNA was eluted in 2 x 20 µl of DEPC-treated water. Vacuum concentration was used to bring the volume of RNA down to 11 µl, ready for amplification.

2.6.4 RNA amplification

After Van Gelder (1990).

cDNA synthesis and RNA amplification were performed using the MessageAmp™ aRNA kit (Ambion (Europe) Ltd, Huntingdon, UK) according to the manufacturers' instructions. Briefly, first round cDNA synthesis was primed with a T7 Oligo (dT) primer. 1 µl T7 oligo (dT) primer was added to 11 µl RNA, heated to 70°C for 10 mins and to this, at 42°C, were added 2 µl 10x first strand buffer, 1 µl ribonuclease inhibitor, 4 µl dNTP mix and 1 µl reverse transcriptase. After a 2 hour incubation, second strand synthesis was performed by the addition of 63 µl nuclease-free water, 10 µl 10x second strand buffer, 4 µl dNTP mix, 2 µl DNA polymerase and 1 µl RNase H. Reactions were incubated for 2 hours at 16°C. cDNA was purified according to the manufacturers' instructions and vacuum concentrated to 16 µl. For *in vitro* transcription, 8 µl of the cDNA was mixed with 8 µl dNTP mix, 2 µl 10x reaction buffer and 2 µl T7 enzyme mix. Reactions were incubated at 37°C for 16-24 hours and then treated to DNase I digestion. aRNA was purified and vacuum concentrated to 10 µl.

Subsequent rounds of cDNA synthesis were performed as follows. 2 µl of random primers was added to 10 µl of aRNA and the reaction was incubated at 70°C for 10 minutes. At 42°C the following were added 2 µl 10x first strand buffer, 1 µl ribonuclease inhibitor, 4 µl dNTP mix and 1 µl reverse transcriptase and the reaction incubated for 2 hours. 1 µl of RNase H was added and incubated for 30 minutes at 37°C. Second strand synthesis was primed with 5 µl T7 Oligo (dT) primer. Following a 10-minute incubation at 70°C the remaining components were added at room temperature: of 58 µl nuclease-free water, 10 µl 10X second strand buffer, 4 µl dNTP mix, 2 µl DNA polymerase. The reaction was incubated at 16°C for 2 hours. cDNA was then purified and *in vitro* transcription performed as described.

Typically three rounds of RNA amplification were required to produce microgram quantities of aRNA.

2.7 cDNA Microarray Analysis

The Affymetrix ATH1 GeneChip® array (approximately 22,800 genes) was used for cDNA microarray analysis. The Arabidopsis Microarray and Bioinformatics service at GARNET performed probe labelling, hybridisation and analysis (Craigon *et al*, 2004). One microgram of cDNA generated from aRNA after three rounds of amplification was provided as the template for probe preparation. For the torpedo root and cotyledon pole tissue, six replicates were performed. For the torpedo shoot apical meristem tissue, three replicates were performed.

2.8 Extraction and purification of nucleic acids

2.8.1 Miniprep of plasmid DNA using the GenElute™ plasmid miniprep kit

Plasmid DNA from high-copy plasmids was isolated using the GenElute™ plasmid miniprep kit from Sigma. The resulting plasmid DNA was suitable for

DNA sequencing and all cloning purposes. Following a modified alkaline-SDS lysis of the cells, plasmid DNA is bound to silica in the presence of high salts. Impurities were then removed by a spin-wash step.

Cells were grown overnight at 37°C in 1-5 ml selective LB media, with vigorous shaking. 1.5 ml of culture was transferred to an Eppendorf tube and the bacterial cells were pelleted by centrifugation at 13,400x g for 1 minute. The supernatant was discarded and the pellet resuspended in 200 µl of resuspension solution containing RNase A to degrade cellular RNA. 200 µl of lysis buffer was added and the tube inverted several times. The lysis reaction was allowed to proceed for 3-4 minutes following which 350 µl of neutralisation/binding solution was added and the tube inverted several times resulting in the formation of a white precipitate containing cellular debris and genomic DNA. The precipitate was pelleted by centrifugation at 13,400x g for 10 minutes. During the centrifugation step a GenElute miniprep binding column was prepared by adding 500 µl of column preparation solution and centrifuging at 12,000x g for 1 minute. The cleared lysate was then transferred to the prepared column and centrifuged at 13,400x g for 1 minute. The flow-through was discarded and 750 µl of diluted wash solution was added to the column and centrifuged at 13,400x g for 1 minute. The flow-through was discarded and the column centrifuged at 13,400x g for 2 minutes to remove residual ethanol. The column was transferred to a fresh Eppendorf and plasmid DNA eluted by the addition of 50 µl of elution solution (10mM Tris-HCL, 1mM EDTA, pH 8) and centrifugation at 13,400xg for 1 minute. The flow-through collected in the Eppendorf contained the purified plasmid DNA.

2.8.2 Midiprep of plasmid DNA using the QIAGEN plasmid midi kit DNA purification system

Plasmid DNA from low-copy plasmids was isolated using the QIAGEN plasmid midi kit DNA purification system (Qiagen, Crawley, England). The resulting plasmid DNA was suitable for DNA sequencing and all cloning purposes. Following alkaline lysis of the cells, plasmid DNA is bound to the QIAGEN Anion-Exchange Resin under low salt and pH conditions. Impurities (including

RNA, proteins and dyes) are removed by a medium-salt wash and plasmid DNA is eluted in a high-salt buffer. Isopropanol precipitation concentrates and desalts the plasmid DNA.

Cells were grown overnight in 100 ml selective LB medium to a density of approximately $4x10^9$ cells/ml and harvested by centrifugation at 6,000 x g for 15 minutes at 4°C. All traces of supernatant were removed and the pellet completely re-suspended in 4 ml Buffer P1 to which RNase A had previously been added. 4 ml of Buffer P2 was added and the contents of the tube mixed by gentle inversion of the tube. This step causes lysis of the cells and care was taken not to shear genomic DNA. Cell debris, including genomic DNA, proteins and SDS was then precipitated by addition of 4 ml Buffer P3, followed by gentle mixing of the tube contents by inversion and incubation on ice for 15-20 minutes. The lysate was cleared by filtration using Miracloth (Calbiochem, La Jolla, California, USA) to remove all suspended and particulate material. Meanwhile a QIAGEN-tip 100 was equilibrated by addition of 4 ml Buffer QBT. The tip works by gravity-flow; solutions are added to the top of the tip, move through the resin, and flow-through is collected at the underneath. The cleared supernatant was added to the equilibrated QIAGEN-tip and allowed to run through before two washes of 10 ml Buffer QC each were used to remove remaining contaminants. Following the second wash, the DNA was eluted with 5ml buffer QF and transferred to 1.5 ml microfuge tubes. Eluted DNA was precipitated by addition of 0.7 volumes room temperature isopropanol, mixed and immediately centrifuged at 15,000x g for 30 minutes at 4°C. The supernatant was washed with 0.4 volumes room temperature 70% (v/v) ethanol and re-centrifuged for 10 minutes at 15,000x g. The ethanol was removed, the pellet air-dried for 5-10 minutes and re-dissolved in 50 µl of 10mM Tris-HCl, pH 8.5.

2.8.3 DNA extraction from plant tissue

To extract DNA from plant cells it is necessary to rupture the cell walls by grinding, and to disrupt the cell membranes by the addition of a detergent. Care must be taken to avoid shearing of the DNA. To prevent endonuclease

digestion of the DNA during extraction, EDTA was included in the extraction buffer as it chelates magnesium ions, which could act as co-factors for endogenous endonucleases. Material from seedlings or young rosette leaves was found to provide the highest yields of DNA.

2.8.3.1 Plant DNA extraction using the GenElute™ plant genomic DNA kit

The GenElute plant genomic DNA kit (Sigma) was used to prepare plant genomic DNA for PCR amplification. Up to 100 mg of plant tissue was ground to a fine powder in liquid nitrogen using a mortar and pestle and transferred to a 1.5 ml Eppendorf tube. Lysis was achieved by the addition of 350 µl of lysis solution (part A) and 50 µl of lysis solution (part B), followed by thorough mixing by vortexing and inversion. The tube was incubated at 65°C for 10 minutes, before 130 µl of precipitation solution was added and the reaction placed on ice for 5 minutes. Cellular debris was pelleted by centrifugation at 13,400x g for 5 minutes. The supernatant was then transferred to a blue filtration column and centrifuged at 13,400x g for 1 minute. DNA was bound by mixing thoroughly by inversion with 700 µl of binding solution. This mixture was transferred to a binding column and centrifuged at 13,400x g for 1 minute. The flow-through was discarded and the column transferred to a fresh collection tube. The column was washed by the addition of 500 µl of wash solution and centrifugation at 13,400x g for 1 minute. The flow-through was discarded and a second wash step carried out. The column was transferred to a new collection tube and the DNA eluted by adding 100 µl elution solution and centrifuging at 13,400x g for 1 minute. The elution step was repeated a second time.

2.8.3.2 Rapid DNA preparation method for PCR

This method was used to isolate genomic DNA rapidly. It yields DNA, which is of lower purity than that from the method described above, but is adequate for the purpose of genomic PCR. Based on Edwards *et al.* (1991).

Solutions

• Extraction buffer: 200mM Tris-HCl pH 7.5, 250mM NaCl, 25mM EDTA, 0.5% (w/v) SDS.

Method

Leaves were collected taking care to avoid cross-contamination and immediately flash frozen in liquid N_2. The samples were then homogenised using a small liquid N_2 -cooled grinder and further liquid N_2 before being thoroughly re-suspended in 400 µl of extraction buffer by vigorous vortexing. The sample was centrifuged at 13,400x g for 5 minutes to pellet cell debris. The cleared supernatant was removed to a new Eppendorf and DNA precipitated with an equal volume of isopropanol. Following centrifugation at 13,400x g for 5 minutes the DNA pellet was washed with 70% (v/v) ethanol. After centrifugation the DNA pellet was allowed to dry and then resuspended in 50 µl sdH_20. 2.5 µl of this DNA prep was used in subsequent 50 µl PCR reactions.

2.8.4 RNA extraction using the Qiagen RNeasy kit

The RNeasy kit from Qiagen was used to prepare total RNA from small amounts of tissue (50-100 mg). Tissue was frozen in liquid nitrogen prior to RNA extraction. The sample was ground under liquid nitrogen to a fine powder in a mortar and pestle and transferred to an Eppendorf containing 450 µl of buffer RLT (10 µl of β-mercaptoethanol added per 1 ml of buffer RLT) and vortexed vigorously. The sample was transferred to the QIAshredder spin column sitting in a 2 ml collection tube and centrifuged for 2 minutes a 13,400x g. The flow-through was removed to a new Eppendorf and 0.5 volumes of ethanol was added and mixed by pipetting. The sample was transferred to an RNeasy mini spin column sitting in a 2 ml collection tube and centrifuged for 15 seconds at 13,400x g. The flow-through was discarded and 700 µl of buffer RW1 was added to the column and the tube centrifuged for 15 seconds at 13,400x g. The flow-through was discarded and the column placed in a new 2 ml collection tube. Two washes were performed with 500 µl buffer RPE by centrifugation at 13,400x g, the first for 15 seconds and the second 2 minutes.

The column was transferred to an Eppendorf and centrifuged at 13,400x g for 1 minute to remove residual buffer RPE. The column was then transferred to a new Eppendorf and 50 µl of RNase-free water was applied to the column followed by centrifugation at 13,400x g for 1 minute to elute the RNA. This step was repeated once more.

2.8.5 Purification of PCR products

PCR products were either run out on an agarose gel and purified using the QIAquick gel extraction kit (Qiagen, Crawley, England) as described in Section 2.8.5.1 or purified directly using the *High Pure* PCR Product Purification kit (Roche, Penzberg, Germany). Nucleic acids bind rapidly and specifically to glass surfaces in the presence of a chaotrophic salt. The bound material can then be separated from salts, proteins and other impurities by washing. Small oligonucleotides, free nucleotides and dimerised primers from PCR do not bind and so are also removed. Nucleic acids elute from the glass surface in a low salt buffer or water. The kits were used according to the manufacturers' instructions.

2.8.5.1 Purification of DNA from agarose gels using the QIAquick gel extraction kit

The QIAquick gel extraction kit was used to purify DNA fragments from agarose (Bioline) gels following restriction enzyme digestion or PCR. DNA fragments were separated on 1% agarose gels in 1x TAE buffer. The gel slice containing the fragment of interest was excised from the gel and placed in 3 volumes (w/v) of buffer QG in an Eppendorf and incubated at 50°C for 10 minutes or until the gel slice had dissolved, the sample was vortexed every 2-3 minutes during the incubation to aid the gel dissolving. 1 gel volume of isopropanol was added to the sample and mixed. The sample was then applied to a QIAquick column and centrifuged at ~17,900x g for 1 minute; the flow-through was discarded. 0.5 ml of buffer QG was then added to the QIAquick column and centrifuged at ~17,900x g for 1 minute, the flow-through was discarded. Adding 0.75 ml of

buffer PE, and centrifuging at ~17,900x g for 1 minute washed the column; the flow-through was discarded. The column was centrifuged at ~17,900x g for an additional minute to remove residual ethanol. The column was then placed into a clean 1.5 ml microcentrifuge tube. DNA was eluted by adding 50 µl of buffer EB to the centre of the QIAquick membrane and centrifuging the column at ~17,900x g for 1 minute.

2.8.5.2 Purification of DNA using the *High Pure* PCR Product Purification kit

The *High Pure* PCR Product Purification kit was obtained from Roche and was used to purify DNA from PCR reactions and restriction enzyme digestion. It was also used to remove oligonucleotide primers (less than 100 bp) from DNA. The DNA containing solution was mixed with binding buffer (100 µl buffer per 20 µl DNA solution) and pipetted into a *High Pure* filter tube in a 2 ml collection tube and centrifuged at 13,400x g for 1 minute. The flow-through was discarded and 500 µl wash buffer added to the filter tube followed by centrifugation at 13,400x g for 1 minute. The wash step was repeated but with 200 µl of wash buffer. The filter tube was placed in an Eppendorf and 50 µl of elution buffer added. DNA was eluted by centrifugation at 13,400x g for 1 minute.

2.9 Electrophoresis

After Sambrook *et al.* (1989).

DNA can be separated, purified and quantified by electrophoresis through an agarose gel matrix. Migration of linear nucleic acid molecules through the gel is inversely proportional to the length of the molecule. Therefore, by loading markers of known size and quantity alongside the samples it was possible to determine their sizes and estimate their amount.

The position of DNA within the gel following electrophoresis can be visualised through the addition of ethidium bromide to the gel. This dye intercalates

between bases of DNA and dye:DNA complexes can be visualised due to their fluorescence under UV light.

Solutions

- 1x TAE buffer: 40mM Tris-acetate pH 8.0, 1mM EDTA
- 10x loading buffer: 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF, 0.25% acridine orange (w/v), and 25% Ficoll (type 400) in water.
- DNA markers: Hyperladder I and Hyperladder IV (Bioline) were used according to the manufacturers' instructions.

Method

1% (w/v) agarose (Bioline) was dissolved in 1x TAE buffer. Gels were melted in a microwave, allowed to cool to approximately 50°C before 1 µl/10 ml of ethidium bromide was added and mixed. The molten agarose was immediately poured into a gel tray and allowed to set at room temperature.

DNA samples were mixed with 1/10 volume of 10x loading buffer to aid loading and to allow the extent of DNA migration to be monitored during electrophoresis. The gel was immersed in 1x TAE buffer in a gel tank. DNA samples and DNA markers were loaded into gel wells by pipetting. DNA markers were run alongside sample DNA to enable approximate sizing of fragments. Electrophoresis was performed at 5-10 V/cm in 1x TAE buffer.

Following electrophoresis DNA was visualised on a UV trans-illuminator (Gel Doc 1000 system) and the image captured using the Molecular Analyst® version 2.1.1 software package (Bio-Rad, Hemel Hempstead, UK) and printed on Mitsubishi K65 HM thermal printer paper. If a Southern blot was to be performed a ruler was placed alongside the gel to enable DNA fragment sizes following hybridisation.

2.10 Polymerase Chain Reaction

2.10.1 Standard PCR

The polymerase chain reaction (PCR) results in the amplification of a segment of DNA between two regions of known sequence, catalysed by *Taq* DNA polymerase purified from the bacterium *Thermus aquaticus*. Oligonucleotide primers are designed which are complementary to the regions of known sequence, and the reaction allowed to proceed through several cycles of denaturation, primer annealing and polymerisation.

For standard PCR reactions, *Taq* DNA polymerase was obtained from Bioline and was supplied with Mg^{++} free 10x reaction buffer and 50 mM $MgCl_2$ stock solutions. Oligodeoxynucleotide primers were obtained from MWG-Biotech as lyophilised pellets and resuspended to the desired concentration in sterile, distilled water.

The template for amplification was either genomic DNA, a cloned fragment of DNA in a plasmid, cDNA or a bacterial colony. A standard PCR reaction contained 10-100 ng of DNA sample, 1 µl of each primer (at 50 pmol/µl), 2 µl 50 mM $MgCl_2$ (2 mM final concentration), 5 µl Mg^{++} free 10x reaction buffer, 1 µl of 25mM dNTP mix and 3.75 units of *Taq* DNA polymerase made up to 50 µl with sterile, distilled water, in a 0.5 ml Eppendorf. The *Taq* Polymerase was always added last. The components were mixed and overlayed with a drop of mineral oil to prevent evaporation during incubation. Reactions were carried out in a DNA Thermal Cycler (Perkin Elmer, Foster City, CA, USA) once the block temperature had reached 90°C.

A typical amplification was carried out using the following conditions: 1 cycle of 5 minutes denaturation at 94°C to ensure complete denaturation of the DNA template, followed by 30 cycles of 1 minute denaturation at 94°C, 1 minute annealing at 55°C and 3 minutes extension at 72°C. A final extension of 7 minutes at 72°C was performed after the amplification steps. The extension step was carried out following the general rule of; 1 minute extension per kilobase of target. PCR products were sized by electrophoresis on a 1% (w/v)

agarose gel alongside known size DNA. Cloning of PCR fragments was carried out as described in Section 2.11.

2.10.2 PCR using Expand™ High Fidelity PCR system

The Expand™ High Fidelity PCR system (Roche) consists of a mix of both *Taq* and *Tgo* DNA polymerases. Due to the 3'-5' exonuclease proofreading activity of *Tgo* DNA polymerase the Expand™ High Fidelity PCR system results in a 3-fold increase in the fidelity of DNA synthesis (4.8×10^{-6} error rate). This system was therefore used when a high degree of sequence fidelity was required, for example, the cloning of DNA fragments for subsequent transformation into plants.

A typical reaction contained 10-100 ng of DNA target, 1 µl of each primer (at 50 pmol/µl), 5 µl 10x reaction buffer, 2 mM $MgCl_2$, 500 µM dNTP mix and 2.5 units enzyme mix made up to 50 µl with sterile, distilled water in a 0.5 ml Eppendorf and overlayed with a drop of mineral oil to prevent evaporation during incubation.

The tubes were placed in a pre-heated block at 90°C. A typical amplification was carried out using the following conditions: denaturation at 94°C for 5 minutes then 30 cycles of denaturation at 94°C for 1 minute, primer annealing at 55°C for 1 minute and extension at 72°C for 3 minutes. This was followed by a final extension at 72°C for 10 minutes. The extension step was carried out following the general rule of; 1 minute extension per kilobase of target. 5 µl of product was then analysed on a 1% (w/v) agarose gel. Cloning of PCR fragments was carried out as described in Section 2.11.

2.10.3 RT-PCR

RNA for RT-PCR was obtained from LCM tissue put through 3 rounds of linear amplification as described in Section 2.6.4. RT-PCR primers were designed around introns to ensure mRNA specificity. Primers were initially tested on total

RNA extracted from 7dpg seedlings using Qiagen's RNeasy mini kit (as described in Section 2.8.4).

The reverse transcription was performed as follows: 5 µg (or 13.7 µl) of RNA was combined in an Eppendorf tube with 1 µl 500 ng/µl Random primers (or oligo $(dT)_{12-18}$ primer in the case of seedling total RNA) and 0.5 µl of 25 mM dNTPs, with the volume made up to 13.7 µl with Rnase-free water. The contents were then mixed and heated at 65°C for 5 minutes. The tube was then transferred to ice and chilled for a further 5 minutes. The contents were spun down briefly in a microcentrifuge and 6.3 µl of the following cocktail added: 4 µl 5x 1^{st} strand buffer, 1 µl 0.1M DTT, 0.3 µl RNasin (40 U/µl) and 1 µl Superscript III (200 U/µl; Invitrogen). The contents were then mixed by pipetting (or gentle vortexing) and spun down briefly. If random primers were used the sample was then incubated at 50°C for 5 minutes. A universal incubation step of 50°C for 2 hours was then performed followed by 15 minutes at 70°C. The samples were then spun down and stored at -20°C before use.

PCR was performed using mRNA specific primers and *taq* polymerase as described in Section 2.10.1.

2.10.4 Genotyping of SALK T-DNA Insertion Lines.

Genotyping of SALK lines required two separate PCR reactions. The first, a 'wild-type' reaction, was designed to amplify a fragment from the gene of interest straddling the T-DNA insertion site. This reaction will only produce a product of the correct size if at least one wild-type copy of the gene is present. A second reaction used a primer specific to the left border of the T-DNA and the reverse wild-type primer. A product should only be produced if the T-DNA was present. An overview of this procedure is given in Figure 2.3. DNA was extracted from plants as described in Section 2.8.3.2 and the PCR carried out as described in Section 2.10.1. No experimental quantification of T-DNA insert number was undertaken.

Figure 2.3 The genotyping of T-DNA insertion lines

Genotyping of SALK lines required two separate PCR reactions:

1. The first, a 'wild-type' reaction, was designed to amplify a fragment from the gene of interest straddling the T-DNA insertion site. This reaction will only produce a product of the correct size if at least one wild-type copy of the gene is present.

2. The second, an 'insert' reaction used a primer specific to the left border of the T-DNA and the reverse wild-type primer. A product should only be produced if the T-DNA was present.

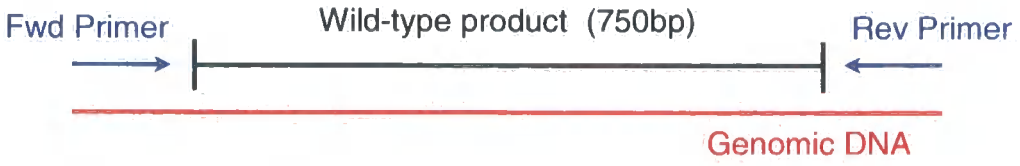
There are three potential outcomes from these PCR reactions, which identify the genotype of the line under investigation:

A. Wild-type – a product is produced by the 'wild-type' reaction only.

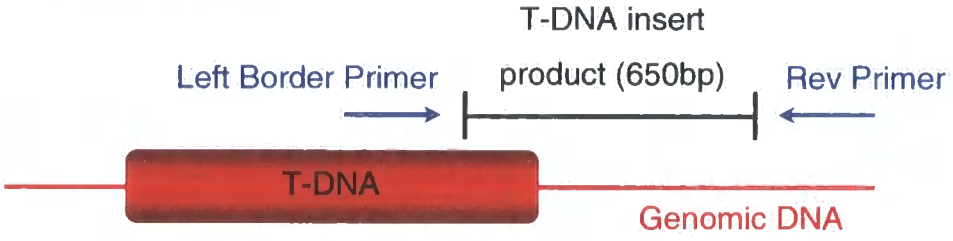
B. Heterozygote – products are produced by both reactions.

C. Homozygote for T-DNA insert – a product is produced by the 'insert' reaction only.

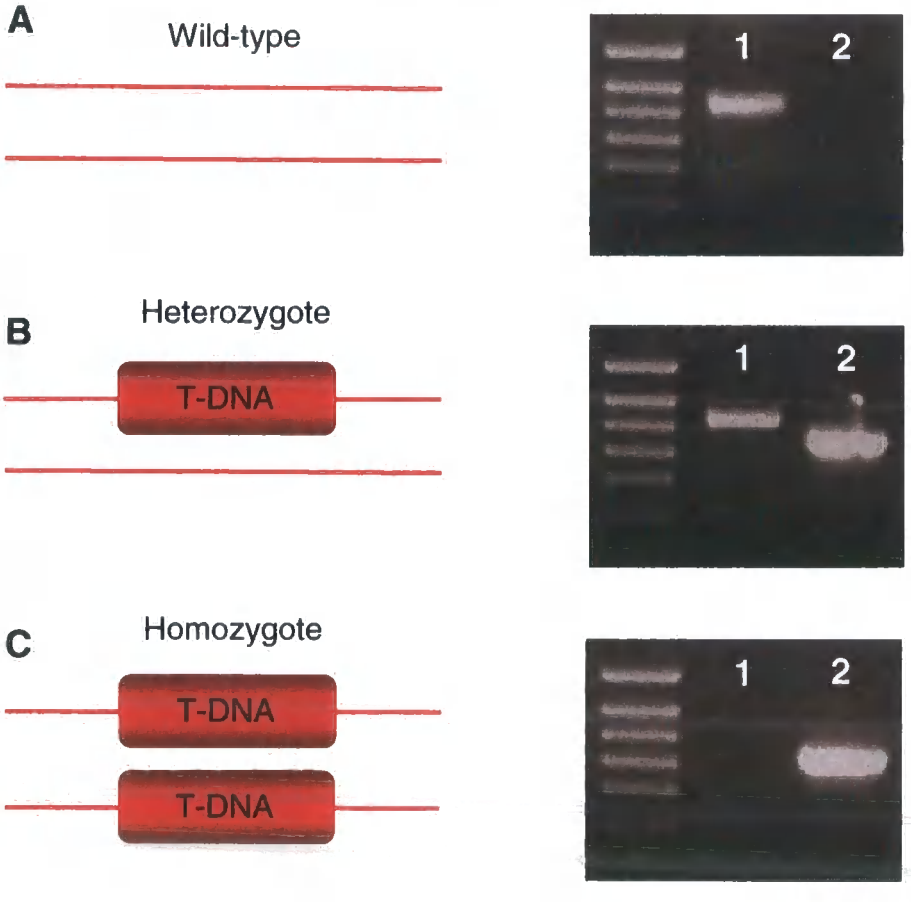
1. Wild-type Reaction



2. T-DNA insert Reaction



3. Potential PCR Genotyping Outcomes



2.11 DNA cloning into plasmid vectors

2.11.1 Digestion of vector and insert DNA with restriction endonucleases

Restriction enzymes (restriction endonucleases) are isolated from a variety of bacteria and are widely used to cleave DNA molecules, both plasmid and genomic. They recognise short palindromic DNA sequences, typically of 4, 5 or 6 nucleotides, at which they cleave the molecule. Restriction enzymes have specific requirements for optimal reaction conditions, being particularly sensitive to the concentration of certain ions. Restriction enzymes and optimised 10x reaction buffers were obtained from Fermentas (Burlington, Canada) or Promega (Southampton, UK). Reactions were carried out according to the manufacturers' instructions.

Typically, a digestion reaction contained 1-5 µg of DNA, 2.5 µl of 10x reaction buffer, 1 µl restriction enzyme 1 (10units/µl), 1 µl restriction enzyme 2 (10units/µl) (if required), and made up to 25 µl with sterile, distilled water. Reactions were left at the required temperature for a period of 1 - 4 hours and terminated by 10 minutes incubation at 70°C.

Following digestion reactions vector DNA was dephosphorylated prior to ligation. Insert DNA was either purified by using the *High Pure* PCR Product Purification kit (Roche) or from gel slices using the QIAquick gel extraction kit.

2.11.2 Addition of 3' Adenine overhangs to digested insert DNA

Restriction enzyme digested insert DNA products lack the 3' adenine residues required for use in TA overhang-dependant cloning reactions, and therefore require the addition of these bases. Initially the digestion product was purified from an agarose gel using the QIAquick gel extraction kit (Qiagen, Crawley, England) and eluted in a volume of 40µl of elution buffer. This was then added to a reaction containing 5 µl of 10x *Taq* PCR Buffer (Bioline), 2 mM $MgCl_2$, 1 µl of 25 mM dNTPs and 1 µl (5 units) of *Taq* (Bioline). The volume was made up

to 50 µl by the addition of sterile, distilled water and incubated at 72°C for 30 min. The enzyme was removed by the subsequent use of a *High Pure* PCR Product Purification kit (Roche).

2.11.3 Dephosphorylation of vector DNA

Following restriction digestion, vector DNA was treated with shrimp alkaline phosphatase (Promega) to dephosphorylate the 5' ends to prevent re-circularisation of the vector and ligation between vector DNA molecules. This was only performed if the vector DNA had been treated with just one restriction enzyme. After digestion, to 1-5 µg of vector DNA with the relevant restriction enzyme in a total volume of 25 µl was added 1 µl (1 unit) of shrimp alkaline phosphatase. The reaction was incubated at 37°C for 1 hour followed by 15 minutes at 70°C to inactivate the phosphatase. Linearised vector DNA was purified from agarose gel slices as detailed in Section 2.8.5.1.

2.11.4 T-tailing of vector DNA

T-tailing of the vector involves the addition of a thymidine nucleotide to the 3' end of DNA strands following vector linearization. This facilitates the cloning of PCR products since *Taq* DNA polymerase adds a 5' adenosine nucleotide. pΔGUS-CIRCE (5 µg) was cut with the restriction enzyme *Sma I* and purified from an agarose gel slice as detailed in Section 2.8.5.1. The linearised vector DNA (40 µl) was mixed with 5 µl Mg^{++} free 10x PCR buffer (Bioline), 2 µl 50 mM $MgCl_2$, 1 µl 25 mM dTTP, 1 µl *Taq* DNA polymerase (5 units) and made up to 50 µl with sterile, distilled water. The reaction was incubated at 72°C for 2 hours and was used in ligation reactions without further purification.

2.11.5 Ligation of DNA fragments

The enzyme T4 DNA ligase joins DNA fragments together by catalysing the formation of a covalent phosphodiester bond between a 5'-phosphoryl group

and an adjacent 3'-hydroxyl group. T4 ligase and 10x ligation buffer were obtained from Fermentas and reactions were carried according to the manufacturers' instruction. In a typical ligation reaction 50-100 ng of vector DNA (cut with a suitable restriction enzyme(s) or T-tailed) was mixed with an equal molar amount of insert DNA. To this was added 2 µl of 10x ligation buffer and 1 µl (3 units) of T4 DNA ligase and the volume made up to 20 µl with sterile, distilled water. The contents were mixed and incubated at 14°C overnight before transformation into competent *E.coli* cells.

2.11.6 Transformation of competent *E.coli* with plasmid DNA.

2.11.6.1 Rubidium Chloride method for preparation of Chemo-competent *E.coli*

Solutions

- Buffer TfbI: 30 mM Potassium acetate, 100 mM Rubidium chloride, 10 mM Calcium chloride, 50 mM Manganese chloride, 15% Glycerol (v/v), pH 5.8.
- Buffer TfbII: 10 mM MOPS, 75 mM Calcium chloride, 10 mM Rubidium chloride, 15% Glycerol (v/v), pH 6.5.

Method

1 ml of an *E.coli* (DH5α) overnight culture were used to inoculate 100 ml of LB broth; this was incubated at 37°C with shaking at 200 rpm to ensure aeration, until an optical density (OD) at 550 nm of 0.5 was achieved. The culture was incubated on ice for 15 min before being pelleted by centrifugation at 3-5000g for 10 min at 4°C. The supernatant was discarded and the pellet gently re-suspended in 40 ml of Buffer TfbI. The cells were incubated on ice for a further 15 min and were subsequently pelleted as before. The pellet was gently re-suspended in 4 ml of Buffer TfbII before being incubated on ice for another 15 min. The cells were then aliquoted into pre-chilled micro-centrifuge tubes and either used immediately or snap frozen in liquid nitrogen and stored at –80°C.

2.11.6.2 Transformation of Chemo-competent E.coli

50 µl of Chemo-competent cells were thawed on ice prior to the addition of 5 µl of a ligation reaction. The cell/DNA mixture was incubated for 30 min on ice before heat shocking at 42°C for 30 seconds. The tube was immediately returned to ice and 500 µl of room temperature SOC medium or LB broth was added. The tube was then incubated for 1 hour at 37°C with gentle shaking. This permits recovery and allows expression of any antibiotic resistance genes encoded by the plasmid to begin. The cells were then spread onto selective plates (50 µg/ml kanamycin) and grown overnight grow at 37°C.

2.11.7 DNA cloning into a plasmid vector using the TOPO TA cloning® kit

PCR products amplified using *Taq* polymerase (Section 2.10.1) may be directly cloned into a plasmid vector using the TOPO TA Cloning® kit (Invitrogen, Paisley, UK). *Taq* polymerase has a non-template-dependent terminal transferase activity that adds a single deoxyadenosine (A) to the 3' ends of PCR products. The linearised plasmid vector (pCR®2.1-TOPO) supplied in the kit has single overhanging, 3' deoxythymidine (T) residues. This allows PCR inserts to ligate efficiently with the vector. TOPO® cloning exploits the ligation activity of topoisomerase I by providing an 'activated', linearised TA vector. Ligation of the vector with a PCR product containing 3' overhangs is very efficient and occurs spontaneously within 5 minutes at room temperature. Cloning and subsequent transformation was carried out according to the manufacturers' instructions.

2.11.7.1 Addition of 3' Adenine Overhangs to PCR Products

PCR products generated by DNA polymerases that lack the ability to add template independent 3' adenine residues (such as *Tgo* the proofreading DNA polymerase used in the Expand high fidelity PCR system) require the addition of these bases before they could be used in TA overhang-dependant cloning reactions. Initially the PCR product was purified using a *High Pure* PCR Product

Purification kit (Roche) and eluted in a volume of 30 µl of elution buffer. This was then added to a reaction containing 5 µl of 10x Taq PCR Buffer (Bioline), 2 mM Magnesium Chloride, 25 mM dATP, and 5 units of *Taq* (Bioline). The volume was made up to 50 µl by the addition of sterile, distilled water and incubated at 72°C for 30 min. The enzyme was removed by the subsequent use of a *High Pure* PCR Product Purification kit (Roche).

2.11.7.2 Ligation of PCR fragments into pCR®2.1-TOPO

DNA fragments generated by PCR were generally cloned into the pCR®2.1-TOPO vector from Invitrogen. To 1 µl of the pCR®2.1-TOPO vector was added 1-2 µl of fresh, unpurified PCR product and 1µl salt solution, the reaction made up to 6 µl with sterile, distilled water. The reactants were mixed gently and left to incubate at room temperature for 5 minutes to allow ligation to proceed. The tube was then placed on ice until ready for transformation into TOP10 competent cells (Invitrogen).

2.11.7.3 Transformation of TOP10 One Shot™ competent cells

TOP10 One Shot™ competent cells were supplied with the TOPO TA Cloning® kit (Invitrogen) along with the ligation ready pCR®2.1-TOPO vector. Following the ligation of PCR products a tube of TOP10 One Shot™ competent cells was defrosted on ice. 2 µl of the ligation mix was added to the cells and mixed by gentle stirring. The tube was incubated on ice for 30 minutes before heat shocking at 42°C for 30 seconds. The tube was immediately returned to ice and 250 µl of room temperature SOC medium was added. The tube was then incubated for 1 hour at 37°C with gentle shaking. 40 µl of 40 mg/ml X-gal was spread onto LB plates containing 50 µg/ml kanamycin sulphate during the incubation. 50-100 µl of cells was then spread onto the selective plates and grown overnight grow at 37°C. Due to disruption of the *lacZ* gene, recombinants would appear as white colonies instead of blue for non-recombinants.

2.12 DNA sequencing

DNA sequencing reactions were performed by the staff of DBS Genomics (Durham) using an ABI 373 DNA sequencer and dye terminator labelling reactions (Perkin Elmer Applied Biosystems). Samples were normally supplied in plasmid form prepared as described in Sections 2.8.1 and 2.8.2. Primers for sequencing were supplied at a concentration of 3.2 pmol/µl.

2.13 Plant transformation

2.13.1 Transformation of *Agrobacterium tumefaciens* with plasmid DNA.

After Shaw (1995)

Solutions

1 mM HEPES/KOH, pH7.0 (filter sterilised from 1 M stock)

20% Glycerol (filter sterilised)

SOC media

Method

0.2 ml of an *Agrobacterium tumefaciens* (strain C58) 2 day culture (grown with antibiotic selection) were used to inoculate 200 ml of selective LB broth (20 µg/ml nalidixic acid and 100 µg/ml streptomycin); this was incubated overnight at 30°C with shaking at 200 rpm to ensure aeration.

The overnight culture of *Agrobacterium tumefaciens* (strain C58) was centrifuged for 10 minutes at 12,000 rpm in a cold (4°C) rotor to pellet. The pellet was re-suspended in 1/3 initial volume of ice-cold 1 mM HEPES/KOH pH 7.0 and re-centrifuged. This was repeated twice before the pellets were re-suspended in ~1/75th initial volume of ice-cold 20% glycerol. The bacteria were then made into 100 µl aliquots in cold Eppendorfs and flash frozen in liquid nitrogen. For electroporation, a 100 µl cell aliquot was defrosted on ice and pipetted into an ice-cold electroporation cuvette (0.2 cm electrode gap, Bio-Rad). 2 µl of supercoiled plasmid from a plasmid prep. (See Section 2.8.2) was

pipetted into the cells and mixed by gentle swirling. Electroporation was carried out using Gene Pulser and Pulser Controller apparatus from Bio-Rad. The Gene Pulser was set to 2.5 kV and 125 μF and the Pulser Controller to 400 Ω. Immediately after electroporation, 1 ml SOC medium was added and the cells were transferred to an Eppendorf tube and incubated for 4-6 hours at 30°C before 100-250 μl aliquots of cells were plated out on selective LB plates (20 μg/ml nalidixic acid and 100 μg/ml streptomycin and 50 μg/ml kanamycin).

2.13.2 *Agrobacterium*-mediated transformation of *Arabidopsis*, using the floral dip method.

After Clough and Bent (1998)

Solutions and media

- 5% sucrose (w/v), 0.05% Silwett L-77 (v/v) (Lehle Seeds. Round Rock, Texas, USA).
- 1/2 MS10: supplemented with 35 mg/l kanamycin sulphate

Method

Arabidopsis thaliana var. Columbia were grown in soil in 3.5" pots (10-15 plants per pot) with a plastic mesh placed over the soil. Primary bolts were removed and the plants grown for 3-4 weeks until the secondary bolts were approximately 10-15 cm tall and displaying a number of immature, unopened flower buds. 2-3 days prior to dipping open flowers and any young siliques were removed. *Agrobacterium tumefaciens* strain C58C3 was used for all binary vector constructs. The *Agrobacterium* were grown for 48 hours at 30°C in 200 ml LB supplemented with 25 mg/l nalidixic acid, 100 mg/l streptomycin sulphate and 50 mg/l kanamycin sulphate. The culture was pelleted by centrifugation and re-suspended in 1 litre of a freshly made solution of 5% sucrose. Once re-suspended, Silwett L-77 was added to a final concentration of 0.05% (v/v). Plants were then dipped fully into the solution and gently agitated for 10-15 seconds before removal. Dipped plants were placed in transparent bags to maintain humidity and placed back in the greenhouse in a shaded position overnight. Occasionally a second dipping was repeated 6 days after the first. Following removal from the bags plants were allowed to set seed and dry out in

the greenhouse. Seed was collected from individual pots of plants and allowed to dry for 2 weeks at 25°C. Seed was surface sterilised and germinated on 1/2MS10 with antibiotic selection. Antibiotic resistant plants were transferred to soil and seed from these plants was tested for segregation on selective plates.

Chapter 3

Results:

Laser Capture Microdissection

3.0 Developing a protocol for the application of Laser Capture Microdissection to embryos of *Arabidopsis thaliana*

3.1 Introduction and objectives

Laser capture microdissection is a relatively recent technological advance, originally developed for the isolation of selected human cell populations from histological sections of complex, heterogeneous tissue (Emmert-Buck *et al.*, 1996). Subsequently this technology has been applied successfully to a variety of animal systems.

The initial objective for this thesis was to develop a protocol for the application of this novel technology to plant tissues, specifically embryos of the model plant species *Arabidopsis thaliana*. Since the onset of this work, other groups have reported successful application of laser capture microdissection to several plant species, including rice, maize and *Arabidopsis* (Matsunaga *et al.*, 1999; Asano *et al.*, 2002; Kerk *et al.*, 2003; Nakazono *et al.*, 2003; Liu *et al.*, 2004; Inada and Wildermuth, 2005), and we have published some of our results (Casson *et al.*, 2005).

The work described by Casson *et al.* (2005) and in this chapter, demonstrates the development of a protocol for the application of LCM to plant embryonic tissue. The objective was to obtain tissue samples that would allow RNA extraction and amplification, for gene expression analyses in discrete regions of the embryo.

3.2 Sample Preparation

Of critical importance in this study was achieving a balance between acceptable histology of tissue sections, thus allowing capture of specific cells, and the preservation of mRNA integrity for subsequent applications.

Excellent preservation of plant tissue morphology is achieved through the use of fixation and embedding protocols. Unfortunately in the context of this work, detrimental effects on the quality of the RNA isolated have been demonstrated to arise from the use of such fixatives and the tissue manipulations involved in the embedding process (Goldsworthy *et al.*, 1999; Gillespie *et al.*, 2002). Goldsworthy *et al.* (1999) demonstrated that fresh frozen mouse liver tissue produced a higher yield of RT-PCR product after laser-capture microdissection than did paraffin-embedded tissue.

An approach utilising fresh frozen plant tissue and cryosectioning was taken, as this was judged to be likely to produce a sufficient yield and quality of RNA from LCM to permit the subsequent applications desired. A potential difficulty that has been observed in mature plant tissue is the formation of ice crystals in the vacuoles and air spaces between cells (Nakazono *et al.*, 2003). Given this possibility, attention was directed to determining the histological integrity of cryosectioned tissues during the LCM process.

It was found that good tissue section histology was achieved by dissecting embryo sacs out of their siliques, embedding in inert OCT medium contained in a plastic mould, freezing in liquid nitrogen-cooled isopentane and then sectioning blocks on a cryostat cooled to -22°C (see Section 2.6.1). No fixation techniques were found to be required for embryonic tissue prior to freezing in isopentane to achieve sections with sufficiently good histology to permit LCM (See Figure 3.1).

Plain uncoated glass slides were found to give the optimal level of tissue adherence for LCM, with sections of between 6-8 μm giving the best transfer. Sections were mounted on slides and fixed immediately in 70% ethanol.

Prior to commencing LCM, tissue sections were serially dehydrated through an ethanol gradient and finally in xylene (See Section 2.6.1). The dehydration and xylene steps are of critical importance to successful LCM, as any moisture present in the tissue sample will prevent proper wetting of the transfer film and consequently impede transfer efficacy

(See http://www.arctur.com/research_portal/support/faq/faq_pixcelliie.htm).

3.3 Laser Capture Microdissection

An initial aim of this work had been to isolate specific, individual embryonic cells and thus be able to dissect very specific tissue types. However, the PixCell II™ LCM system used was slightly limited in this respect in that the minimum laser beam diameter available was 7.5 μm. Even with this laser beam diameter it was found to be impossible to accurately dissect out specific, single embryonic cells, which were often of a smaller diameter than this. Therefore, the system was found to be limited to capturing small clusters of cells in a reproducible fashion from a predefined area (15 μm for the cotyledon and root poles and 7.5 μm for the shoot apical meristem). Despite testing a number of different slide coating treatments, it was not found to be possible to reproducibly achieve complete transfer of the small clusters targeted, therefore the targeting of a specific cluster was repeated a number of times to improve the transfer. In addition, replicates of cell types of interest were targeted from multiple embryo samples obtained from different siliques of an individual plant and from different individual plants to account for any potential biological variation. Approximately 15 such replicate cell clusters were captured per CapSure™HS LCM cap, ready for RNA isolation.

To support and further the work described by Casson *et al.* (2005) on globular and heart stage embryos of *Arabidopsis thaliana*, torpedo stage embryos were chosen for analysis. The root pole and cotyledonary tissues were sampled in addition to the presumptive shoot apical meristem. The results presented in Figure 3.1 show that, using the sample preparation techniques described above and in Sections 2.6.1 and 2.6.2, the capture of all three tissue samples was successful, with little obvious contamination with other tissues.

Figure 3.1 Laser capture microdissection of cryosections of torpedo-stage *Arabidopsis* embryos.

(a-j) LCM of torpedo-stage embryo.

(a) Torpedo-stage embryo.

(b) Root pole after targeting with laser.

(c) Root pole after removal of the cap.

(d) Root pole cells on the cap.

(e) Cotyledon pole after targeting with laser.

(f) Cotyledon pole after removal of the cap.

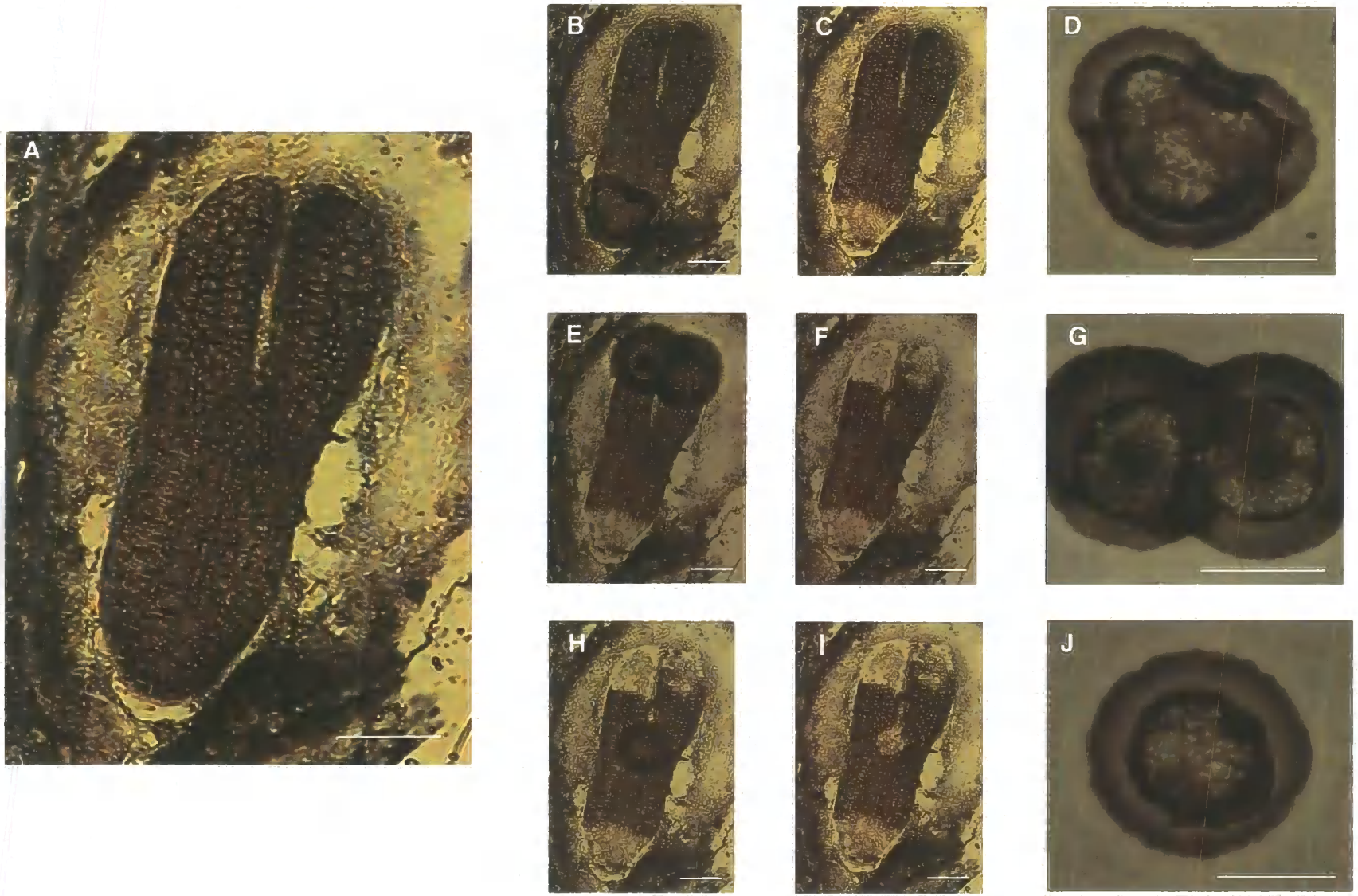
(g) Cotyledon pole cells on the cap.

(h) Shoot apical meristem after targeting with laser.

(i) Shoot apical meristem after removal of the cap.

(j) Shoot apical meristem cells on the cap.

Scale bars in A-G represent 15 μm, scale bar in F represents 7.5 μm.


A
B
C
D
E
F
G
H
I
J

3.4 RNA amplification

As Zimmerman and Goldberg (1977) estimated that a single tobacco leaf cell contains approximately 38 pg of polysomal RNA, and LCM on embryonic tissue does not necessarily capture a complete cell, then the expected yield of RNA from captured cells would be insufficient for either accurate quantification or subsequent GeneChip® analysis.

In light of the limited yield of RNA expected from LCM embryonic cells, and the conclusion that obtaining the amount of embryonic tissue required for the isolation of µg quantities of RNA was not feasible, an RNA amplification step was required to generate sufficient RNA for downstream applications.

Maintaining the population distribution of transcripts is a major problem associated with the amplification of such small quantities of mRNA. The amplification could not be accomplished using standard PCR methods as these are generally considered to produce a bias due to the preferential amplification of the smaller and more abundant transcripts in an mRNA population. Therefore a linear T7 RNA polymerase mediated amplification technique was used, as developed by Van Gelder *et al.* (1990), and which has previously been utilised to amplify RNA derived from animal tissue prior to microarray analysis.

Following DNase treatment, the RNA isolated from LCM cells was subjected to three rounds of amplification as described in Section 2.6.4. This generated approximately 5-10 µg of amplified RNA (aRNA) as determined by spectrophotometric measurement. Gel analysis indicated a size distribution of between 100 and 1000 nt (Figure 3.2).

3.5 RT-PCR analysis

To ensure that the aRNA generated was a suitable template for analysis and also to assess prior to microarray analysis if differential gene expression could be detected in different tissue samples, RT-PCR was performed (see Section 2.10.3). Genes were selected for this analysis based on distinct apical-basal

Figure 3.2 Gel analysis of aRNA from torpedo embryos after three rounds of amplification.

Lane 1: Hyperladder IV (Bioline), 1 Kb – 100 bp

Lane 2: Torpedo cotyledon – Replicate 1

Lane 3: Torpedo root – Replicate 1

Lane 4: Torpedo shoot apical meristem – Replicate 1

Lane 5: Hyperladder I (Bioline), 10 Kb – 100 bp

Lane 6: Torpedo cotyledon – Replicate 2

Lane 7: Torpedo root – Replicate 2

Lane 8: Torpedo shoot apical meristem – Replicate 2

Lane 9: Hyperladder I (Bioline), 10 Kb – 100 bp

Lane 10: Torpedo cotyledon – Replicate 3

Lane 11: Torpedo root – Replicate 3

Lane 12: Torpedo shoot apical meristem – Replicate 3

Lane 13: Hyperladder IV (Bioline), 1 Kb – 100 bp

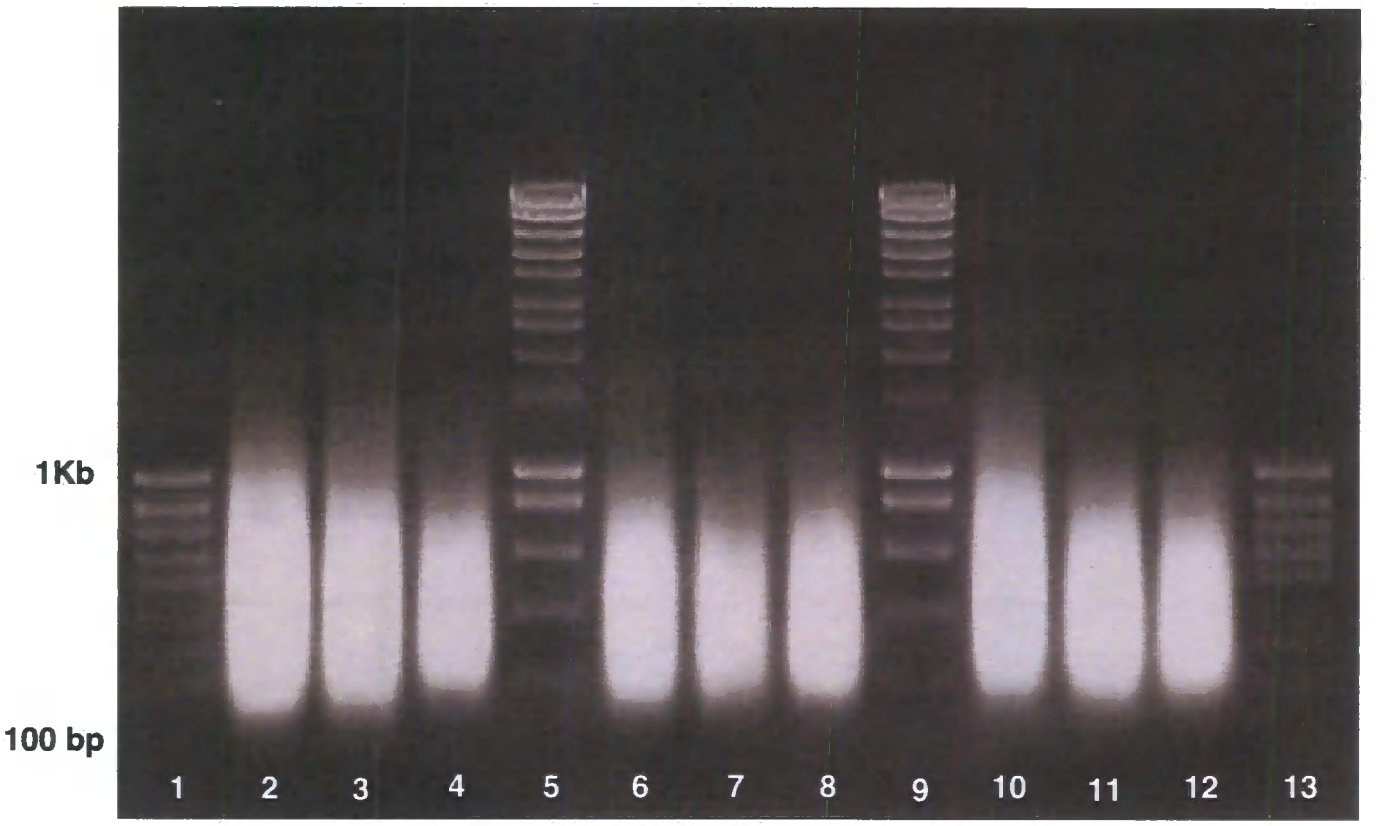

1Kb
100 bp
1
2
3
4
5
6
7
8
9
10
11
12
13

Figure 3.3 RT-PCR analysis of aRNA

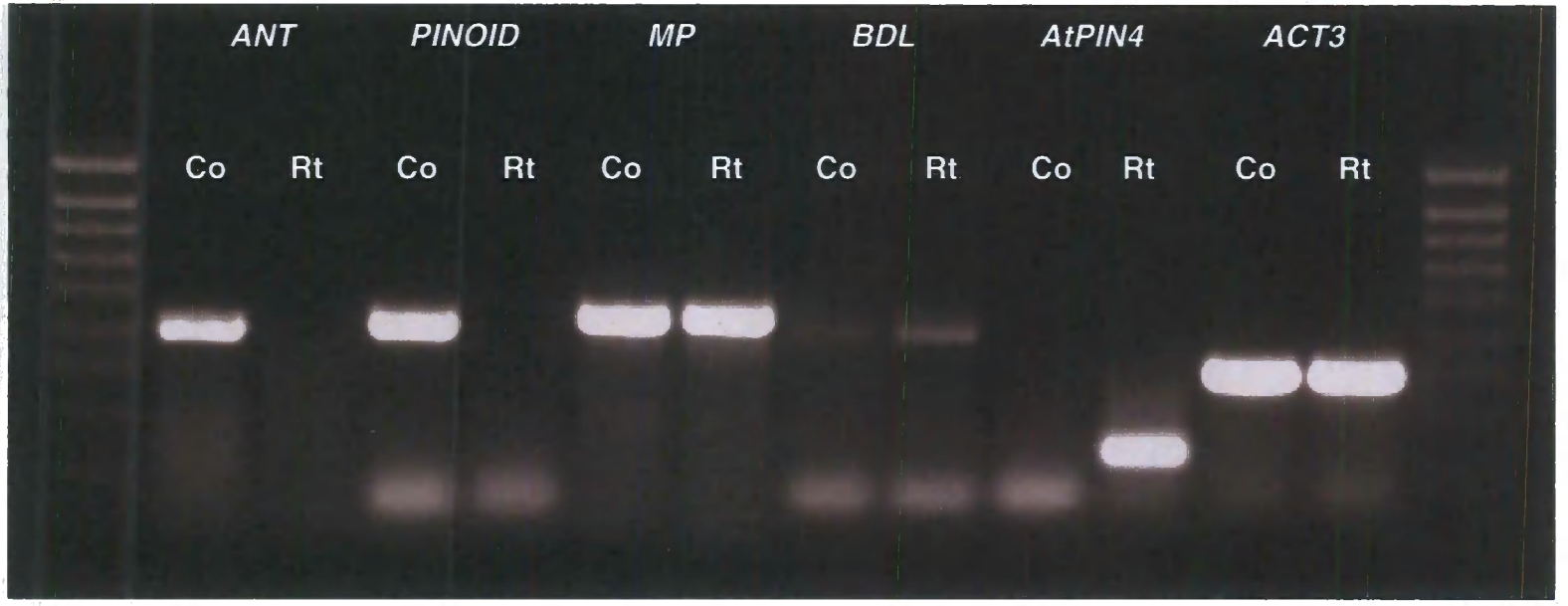
RT-PCR analysis was performed on aRNA from torpedo stage embryos after three rounds of amplification to determine its suitability as a template and as a preliminary validation of the technique to show differential expression of key embryonic genes with published expression patterns.

Lanes 1 and 14: Hyperladder IV (Bioline) 1 Kb – 100 bp

Co – Cotyledonary region
Rt – Root region

ANT - AINTEGUMENTA
PINOID - PINOID
MP - MONOPTEROS
BDL - BODENLOS
AtPIN4 - AtPIN4

ACT3 – Predicted constitutively expressed control gene


ANT
PINOID
MP
BDL
AtPIN4
ACT3
Co
Rt
Co
Rt
Co
Rt
Co
Rt
Co
Rt
Co
Rt

expression patterns published in the literature (see Section for 4.4 for a more comprehensive comparison and references). The results of this analysis (using mRNA specific primers designed around introns), for the genes *AINTEGUMENTA (ANT), PINOID, MONOPTEROS (MP), BODENLOS (BDL)* and *AtPIN4* are shown in Figure 3.3. The results show *AINTEGUMENTA (ANT)* and *PINOID* to be present only in the cotyledon sample, and *AtPIN4* to be present only in the root sample. *MONOPTEROS (MP)* and *BODENLOS (BDL)* were present in both cotyledon and root samples, *MP* displayed a constitutive expression pattern whereas *BL* showed a slightly higher expression in the root. All of these expression patterns corresponded to those published in the literature. The *ACT3* control also showed constitutive expression indicating that there was no major bias in template availability between the two samples.

3.6 Discussion

3.6.1 Sample preparation

The methods used to prepare samples in terms of the fixation and processing parameters employed have been shown by numerous studies to have significant effects on the molecular profile of the tissue sample, and thus are of critical importance in the experimental design.

Fixation aims to stabilise cell contents and preserve the quality and integrity of a specimen during processing, however, this in itself introduces artefacts due to the chemical modifications required (Srinivasan *et al.*, 2002).

The quality and yield of RNA were of principal concern in this study, as a sufficient amount of high quality RNA was deemed an essential prerequisite to obtaining reliable microarray data, and thus informed the choices taken.

Two types of chemical fixatives have been assessed with regard to their effect on plant tissue used in LCM studies. Formaldehyde-based fixatives are known to provide superior preservation of morphology but were shown to result in both poor yields and degradation of nucleic acids (Goldsworthy *et al.*, 1999; Kerk *et*

al., 2003; Nakazono *et al.*, 2003) and were thus avoided. Precipitating fixatives (such as acetone and ethanol) have previously been shown to be preferable in terms of the yield and quality of RNA obtained from animal cells (Goldsworthy *et al.*, 1999). Studies on plant tissue have confirmed this observation (Kerk *et al.*, 2003; Nakazono *et al.*, 2003). A zinc-based fixative has also been demonstrated to recover RNA from animal cells at a similar efficiency to direct homogenisation, although to date this has not been employed in the fixation of plant tissue for LCM (Scheidl *et al.*, 2002).

However, given all the fixation methods available, freezing tissue is still regarded as the best method currently available for subsequent molecular analysis in animal systems. A study of the expression profiles of microdissected tumour cells on Affymetrix microarrays showed that, while frozen samples gave an average present call (Affymetrix MAS5 software uses a detection algorithm based on probe pair intensities compared to a predefined threshold Tau, to calculate a detection p-value, and assign a present, marginal or absent call for each gene) of 26%, ethanol-fixed paraffin-embedded samples gave a present call of only 4.5% thus highlighting previous studies that noted increased yields and quality of RNA from frozen tissue (Goldsworthy *et al.*, 1999; Gillespie *et al.*, 2002; Perlmutter *et al.*, 2004). Fixation, post-sectioning, in 70% ethanol as used in this work, was shown by Goldsworthy *et al.* (1999) to be optimal for the recovery of RNA from laser microdissected tissue.

In a study of tumour samples, OCT-embedded samples were shown to be comparable to fresh frozen tissue samples in terms of the gene expression profile observed from oligonucleotide microarray analysis (Sanchez-Carbayo *et al.*, 2003). Thus, it can be extrapolated that no detrimental effect on RNA integrity resulted from embedding torpedo-stage embryos in this medium prior to freezing.

Kerk *et al.* (2003) investigated the RNA yield returned from tissue sections of differing thickness (3 µm to 10 µm) of radish (*Raphanus sativus)* cortical parenchyma and found little difference in this range. They standardised the section thickness used to 10 µm for mature tissue and 6 µm for developing

tissue with comparatively smaller cells. The 8 µm section thickness routinely employed in this work is comparable to this study.

The formation of ice crystals in vacuoles and air spaces of fresh, frozen plant material has been reported to have detrimental effects on specimen morphology thus prohibiting its use and necessitating the requirement for some form of fixative (Nakazono *et al.*, 2003). However, the loss of histological integrity in fresh, frozen sections of embryos as described in this chapter was deemed not to be significant in terms of target acquisition when compared to the potential losses of RNA integrity that would result from chemical fixation.

That the tissue histology was determined to be sufficient for this work is a consequence of both the small, cytoplasmically dense embryonic cells not being as susceptible to loss of cytological integrity as mature tissue, and of limitations to the resolution of the laser capture microdissection apparatus employed. The minimum beam setting on the Arcturus PixCell II system was only 7.5 µm, which ruled out the capture of single embryonic cells and therefore the histology was only required to be sufficient to identify overall embryo morphology. Using a system capable of microdissecting a smaller area would potentially have necessitated a fixation step to fully exploit the capacity to target individual cells. The yields of RNA obtainable from paraffin-embedded tissue are still acceptable when the focus becomes the superior morphology required for the identification of specific cell types (Kerk *et al.*, 2003; Perlmutter *et al.*, 2004; Inada *et al.*, 2005).

3.6.2 Laser Capture Microdissection

There are a number of laser-mediated microdissection systems available, which fall into two discrete categories based on the method used for target isolation. Laser capture microdissection (LCM) systems, such as the Arcturus PixCell II system used in this work are based on the original system developed by Emmert-Buck *et al.* (1996) and have been used successfully in many applications. Data indicate that the LCM system does not have a deleterious effect on the recovery of RNA from microdissected tissue as any thermal

changes experienced by the tissue are mild and of very-short duration (ms) (http://www.arctur.com/research_portal/support/faq/faq_pixcellIIe.htm). The alternative system is that of laser excision using a narrow ultraviolet laser beam to ablate tissue surrounding the cells of interest, thus cutting out the region of interest. The UV laser is of a higher wavelength than the absorption peak of nucleic acid and is therefore not predicted to cause thermal changes to the adjacent cells (http://www.palm-mikrolaser.com/dsat/index.php). The PALM® MicroLaser system is combined with laser pressure catapulting (LPC) which ejects the isolated sample with a single defocused laser pulse. Here the sample is not in contact with any part of the equipment during the procedure, which dramatically reduces any chance of contamination, which could be present in other systems (Scheidl *et al.*, 2002).

A major advantage of the PALM® system particularly when applied in the context of this work is the potential reduction in the area and shape of the isolated region. Using a beam spot size of <1µm in diameter the system permits the cutting out of cells of all shapes and sizes, down to the resolution of pieces of chromosomes (Scheidl *et al.*, 2002; http://www.palm-mikrolaser.com/dsat/index.php). Such resolution would open up an entirely new range of specificity and would permit a range of different questions to be asked. However as described above, this would also necessitate a requirement for superior histology and thus require a compromise to be made in terms of fixation with its associated loss of RNA yield and integrity.

3.6.3 RNA amplification

RT-PCR analysis of gene expression has been demonstrated directly from RNA recovered from laser microdissected cells (Kerk *et al.*, 2003). However, for downstream applications such as microarray analysis microgram quantities of RNA are required in order to generate sufficient hybridisation signal to allow accurate quantification. Given that LCM generally yields nanogram or picogram quantities of RNA a highly efficient linear amplification step was required.

Standard PCR is generally regarded to produce an un-reproducible bias towards smaller and more abundant transcripts. However, specialised PCR protocols have been developed which attempt to achieve a greater linearity of amplification by limiting product size to such a degree that only the 3' end of the cDNA is amplified and thus all transcripts are amplified equally (Iscove *et al.*, 2002). This method would seem more suited to cDNA microarrays rather than Affymetrix arrays as transcript truncation would result in many redundant probe pairs for each transcript.

The linear amplification mediated by T7 RNA polymerase used in this work has been successfully applied to RNA extracted from laser microdissected animal tissue for subsequent microarray applications, and more recently on plant tissue (Salunga *et al.*, 1999; Kerk *et al.*, 2003).

A number of studies have been undertaken to assess the reliability and reproducibility of expression profiles, using both cDNA and oligonucleotide arrays, to determine the effects of T7 amplification (Scheidl *et al.*, 2002; Nakazono *et al.*, 2003; King *et al.*, 2005). A study using T7 amplified RNA from laser microdissected primary breast tissue determined that any technical variability introduced by the isolation technique and the amplification procedure was small relative to differences in biological variability observed between comparative tissues (King *et al.*, 2005). Nakazono compared T7-amplified RNA from laser microdissected material from maize coleoptiles with a comparable amount (40 ng) of non-amplified RNA using a two-colour cDNA microarray and demonstrated a highly linear relationship demonstrating reproducibility among samples. These studies only assessed changes to the expression profile after 2 rounds of amplification rather than the 3 used here. However the results of Scheidl *et al.* (2002) suggested that further rounds of amplification would produce no significant increase in variability. These authors suggested that the difference in expression profile observed between amplified and non-amplified samples results not from a bias resulting from experimental variability but from the global reduction in transcript length resulting from priming with random hexamers. Nygaard *et al.* (2005) addressed the correlation between quantity of input material and data reliability, and concluded that at extremely low quantities of starting material, such as single cell analyses, stochastic

fluctuations become significant possibly affecting reliability (Nygaard *et al.*, 2005).

Laser microdissection was developed to allow the isolation of a homogenous cell population from a heterogeneous tissue (Emmert-Buck *et al.*, 1996). Therefore in comparison to a global study of heterogeneous tissue it can be predicted that the sampling will reduce the biological variability, whilst technical variability will be added as a result of the amplification protocol. Direct comparison between expression data from amplified and non-amplified samples must take these caveats into account.

3.6.4 Validation of the sample preparation, isolation, extraction and amplification procedures by RT-PCR analysis.

In order to validate the success of the protocols used to obtain sufficient RNA from the laser capture microdissection of torpedo-stage embryos for microarray analysis, RT-PCR was performed on 3rd round aRNA using primers for genes with known expression patterns, namely *AINTEGUMENTA (ANT)*, *PINOID (PID)*, *AtPIN4*, *MONOPTEROS (MP)* and *BODENLOS (BDL)*.

In situ hybridisation analysis show clear localisation to the cotyledons of the torpedo-stage embryo of the *ANT* transcript (Elliot *et al.*, 1996), and until the mid torpedo-stage embryo *PID* is localised to the outer layers of the cotyledons (Christensen *et al.*, 2000), both of these expression patterns are clearly seen in the RT-PCR result with expression only detectable in the cotyledon sample.

AtPIN4, a member of the *PIN* family of putative auxin efflux carriers was selected as a basal control, whole-mount *in situ* immunolocalisation studies have detected ATPIN4 in and around the root meristem, a localisation confirmed by *in situ* hybridisation and promoter::GUS studies of the *ATPIN4* expression pattern. The RT-PCR analysis mirrored this expression pattern with a product produced only from the root sample.

The expression patterns of *MP* and *BDL* have been analysed by *in situ* hybridisation and show the two genes to be co-expressed with both mRNA species restricted to the vascular precursor cells by the torpedo-stage (Hardtke and Berleth, 1998; Hamann *et al.*, 2002). The RT-PCR analysis again fits with the published patterns in that a product for both genes is found in both cotyledon and root samples. While the *MP* expression appears to be entirely constitutive, the semi-quantitative RT-PCR indicates that there is more *BDL* expression in the root, this correlates well with a *pBDL::GUS* construct which indicates that *BDL* expression is at its highest in the root, the effect of the *bdl* mutation on primary root formation underlines its importance in this region (Hamann *et al.*, 1999; Hamann *et al.*, 2002).

Given that the RT-PCR analysis confirmed the expression patterns of cotyledon specific, root specific and constitutive genes (including the predicted constitutive *ACT3* control) the protocols employed to reach 3rd round aRNA are validated as sufficient to produce high quality RNA for subsequent microarray analysis.

The following chapter will describe the microarray analysis and a variety of methods used to validate the data obtained.

Chapter 4

Results:

Validation of Microarray Analysis

4.0 Validation of microarray data

4.1 Introduction and objectives

In the previous chapter the application of the novel technology of laser capture microdissection was successfully demonstrated on embryonic tissue of the model plant species *Arabidopsis thaliana*. The RNA was isolated and amplified using a linear T7 RNA polymerase mediated amplification technique.

RT-PCR was performed using genes of known expression pattern, however the amount of potential information gained in this way is limited, its main purpose in this thesis being to validate the extraction and amplification methods. Therefore it was decided to use the Affymetrix ATH1 GeneChip® DNA microarray system (Affymetrix Inc., Santa Clara, USA) to allow a global transcriptional analysis.

The objective of the following work was initially to ascertain whether or not the aRNA from LCM embryonic tissue was suitable for application to a microarray platform. Subsequently, validation of the data produced was of critical importance and this was achieved though a number of methods. Firstly, comparisons were made between GeneChip® expression data for LCM material and published expression patterns for known embryonic genes obtained from *in situ* hybridisation and promoter::GUS analysis. Secondly, through the creation of promoter::GUS constructs for putative transcription factors showing a bias towards either the cotyledon or root selected from the microarray data. Additionally an attempt was made to assign a number of these putative transcription factors a potential functional role through the analysis of T-DNA insertion lines obtained from the SALK institute.

In this way it was hoped to confirm the validity of the use of LCM on embryonic tissue to produce aRNA of sufficient quality for labelling and subsequent hybridisation to a GeneChip®, and further to validate the expression data produced, thus providing a degree of confidence in the data set necessary to begin higher level bioinformatic analysis

4.2 Microarray analysis

Six biological replicates were performed for each tissue (torpedo cotyledon and torpedo root), in addition to three replicates of torpedo shoot apical meristematic tissue. Each biological replicate was obtained from a separate batch of plants grown at a different time in a distinct region of the growth room.

Additional quality control was performed on the aRNA samples by the Arabidopsis microarray and bioinformatics service provider at the Nottingham Arabidopsis Stock Centre (NASC) to ensure it was of suitable quality for microarray applications (NASC uses the Agilent Bioanalyzer to check the integrity of RNA, and a Nanodrop to calculate the concentration and assess the purity of RNA). The aRNA satisfied the quality control inspection and was allowed to proceed through to microarray analysis.

Due to the 3' bias present in aRNA compared to unamplified mRNA, it was necessary to reduce the number of probe pairs used in the GeneChip® analysis, restricting them to those designed towards the 3' end of transcripts. The enrichment of the 3' end of transcripts compared with the 5' occurs during cDNA synthesis, as a consequence of the random priming required during first or second strand synthesis. Each additional round of cDNA synthesis and RNA amplification leads to an increase in the 3'-5' bias.

The ATH1 GeneChip® data produced by this study can be fully accessed through the NASC (Craigon *et al.*, 2004; http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl). Signal and detection call values were generated by Affymetrix Microarray Analysis Suite 5.0 software (Affymetrix Inc., Santa Clara, USA).

4.3 Estimation of the number of genes expressed

The use of the ATH1 GeneChip® which has a coverage of ~22,800 genes allows an estimation to be made of the number of genes which are being expressed in the tissue under analysis, and also a comparison to other

analyses performed on the same microarray platform. The use of 'estimate' refers to the fact that a cut-off point of a minimal Affymetrix signal must be imposed; genes deemed to be expressed are those with a value equal to or greater than this value, those less than the defined signal value are deemed not to be expressed. The application of a minimal Affymetrix signal value of 75 has been applied to a study of the root transcriptome (Birnbaum *et al.*, 2003). Recently, a minimal signal value of 40 was applied to a study of globular and heart stage embryonic tissue (Casson *et al.*, 2005).

To put my data in the context of published literature it was decided to calculate an estimated number of expressed genes for both of these cut-off values. The replicates for each tissue type (torpedo SAM, cotyledon and root) were collated and a mean signal value across the replicates calculated. The mean signal values were ranked highest to lowest thus revealing the number of genes with an equal or greater value than the designated cut-offs. The results are shown in Table 4.1. Using the cut-off value of 75 for the mean value of the replicates analysed, my data indicates that between 8353 and 11,690 genes (~37-51%) are expressed in the three tissue types of the torpedo stage embryo. If the lower mean signal threshold value of 40 is applied up to ~77% of the genes are deemed to be expressed.

A spatial analysis was then performed on genes predicted to be expressed, to ascertain what degree of overlap exists between the different tissue types. This analysis is summarised in Figure 4.1, in the form of a Venn diagram for each cut-off value. The majority of expressed genes are present in all tissue types (~60% when using the signal threshold value of 40 and ~43% when using the signal threshold value of 75%). There are also significant numbers of genes present in single tissue types only (~18% when using the signal threshold value of 40 and ~29% when using the signal threshold value of 75), thus suggesting distinct spatial transcriptional profiles are present.

Table 4.1 Estimate of the number of genes expressed based on a signal value cut-off.

The ATH1 GeneChip® has a coverage of ~22,800 genes allowing an estimation to be made as to the number of genes expressed in the tissue under analysis. The use of 'estimate' refers to the fact that a cut-off point of a minimal Affymetrix signal must be imposed; genes deemed to be expressed are those with a value equal to or greater than this value, those less than the defined signal value are deemed not to be expressed. The application of a minimal Affymetrix signal value of 75 has been applied to a study of the root transcriptome (Birnbaum *et al.*, 2003). Recently, a minimal signal value of 40 was applied to a study of globular and heart stage embryonic tissue (Casson *et al.*, 2005).

An estimated number of expressed genes for both of these cut-off values was calculated for the ATH1 GeneChip® data for torpedo stage embryonic tissue. The replicates for each tissue type (torpedo SAM, cotyledon and root) were collated and a mean signal value across the replicates calculated (these comprised six replicates for the cotyledon and root and three replicates for the SAM). The mean signal values were ranked highest to lowest thus revealing the number of genes with an equal or greater value than the designated cut-offs.

Table 4.1 Estimate of the number of genes expressed based on a signal value cut-off.

Tissue type (all torpedo stage)	No. of genes with a signal value ≥40	No. of genes with a signal value ≥75	Total no. of genes on ATH1 GeneChip®	% Present at signal value cut-off =40	% Present at signal value cut-off =75
SAM	16,023	11,189	22,746	70.44	49.19
Cotyledon	13,454	8353	22,746	59.15	36.72
Root	17,479	11,690	22,746	76.84	51.39

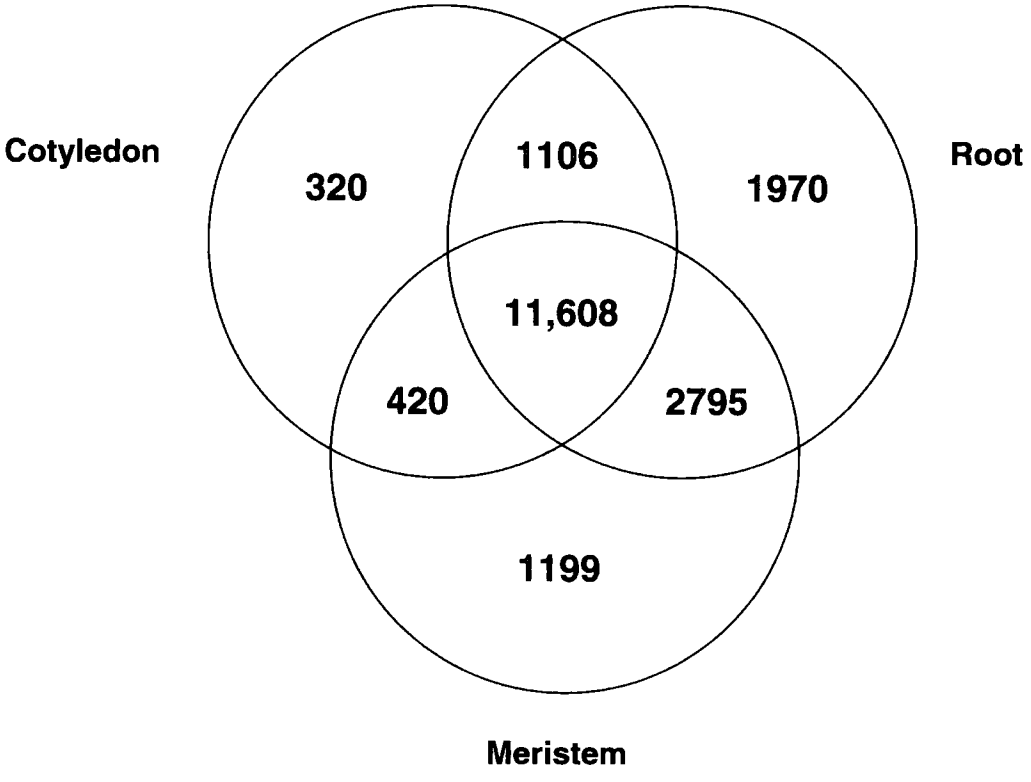
Figure 4.1 Venn diagrams showing overlapping expression of genes between the tissue regions of the torpedo-stage embryo.

An estimated number of expressed genes for both signal cut-off values (40 and 75) were calculated. The replicates for each tissue type (torpedo SAM, cotyledon and root) were collated and a mean signal value across the replicates calculated. The mean signal values were ranked highest to lowest thus revealing the number of genes with an equal or greater value than the designated cut-offs.

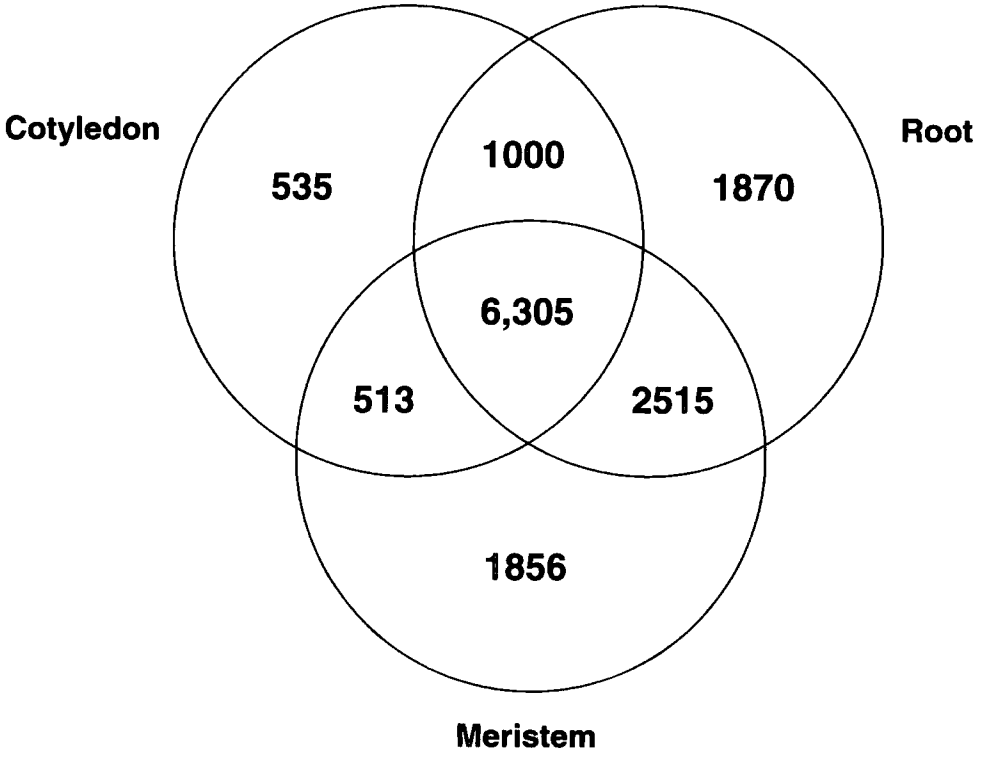
A spatial analysis was then performed on these expressed genes, determined at each signal cut-off value, to ascertain what degree of overlap exists between the different tissue types.

A Venn diagram is presented for each signal cut-off value, with the overlapping regions corresponding to the number of expressed genes present in more than one tissue type. The central region corresponds to the expressed genes present in all tissue types.

Mean signal value >40



Mean signal value >75



4.4 Comparison with published expression patterns of embryonic genes

In order to make deductions from my microarray data set in any way justifiable it is of critical importance to achieve some level of validation. One approach to this problem is to compare the expression data obtained from my LCM samples with expression pattern data previously reported for embryonic genes in the literature. If expression levels from my GeneChip® analysis correspond to expression patterns confirmed in the embryo by other workers through *in situ* hybridisation or promoter::GUS analysis then this would provide an initial level of validation for the approach.

A number of embryonic genes were selected based on well-characterised expression patterns confirmed by *in situ* hybridisation or promoter::GUS analyses; the results of this comparison are shown in Table 4.2. The expression levels from my GeneChip® analysis correspond very well with the previously reported expression patterns reported in the literature, highlighting the known cotyledon expressed genes *AINTEGUMENTA* and *PINOID*, the signal values for both are in excess of 5-fold higher in the cotyledons compared to the roots. The root pole marker genes *AtPIN3* and *AtPIN4* have much lower signal values than those of *AINTEGUMENTA* and *PINOID* but show a distinct elevation in the root tissue compared to the cotyledon. Thus my GeneChip® analysis compares favourably to the preliminary RT-PCR analysis presented in Section 3.5, demonstrating a successful transition from amplified RNA through to microarray output.

4.5 Identification of differentially regulated putative transcription factors for promoter::GUS analysis

The number of previously characterised embryonic genes with distinct spatial expression patterns is relatively low and thus this approach to validating my GeneChip® data is limited. It was therefore decided to further validate the data through the selection of genes of unknown function, which appear from the

Table 4.2 The expression of embryonic genes based on *in situ* or promoter::GUS analysis derived from the literature, compared with their expression determined by GeneChip® analysis of LCM embryonic tissue.

Gene	Expression in torpedo stage embryos	Reference	Torpedo Cotyledon	Torpedo Root
AINTEGUMENTA	Cotyledons	Elliot *et al.*, 1996; Long and Barton, 1998	647.28 (±298.3)	20.73 (±25.2)
ARGONAUTE	Throughout	Lynn *et al.*, 1999	180.16 (±135.5)	299 (±371.9)
AtPIN1	Vascular precursor cells	Steinmann *et al.*, 1999	120.1 (±29.7)	164.83 (±126.2)
AtPIN3	(Root Pole – Heart Stage) (Root Pole – Seedling)	Friml *et al.*, 2003	66.05 (±33.8)	143.2 (±134)
AtPIN4	Root Pole	Friml *et al.*, 2002	75.26 (±73.2)	118.3 (±48.8)
BODENLOS	Vascular precursor cells	Hamann *et al.*, 2002	49.04 (±25.5)	56.75 (±33.1)

FUS3	Protoderm	Tsuchiya *et al.*, 2004	198.5 (±90.7)	213.9 (±233.6)
LEC1	Outer protoderm and ground tissue	Lotan *et al.*, 1998	540.68 (±291)	222.6 (±226.3)
MONOPTEROS	Vascular precursor cells	Hardtke and Berleth, 1998 Hamann *et al.*, 2002	109.32 (±44)	95.8 (±109)
PINOID	Outer layers of cotyledon	Christensen *et al.*, 2000	505.78 (±454.3)	75.93 (±101.7)
ZWILLE	Vascular precursor cells	Moussian *et al.*, 1998	410.08 (±219.3)	505.05 (±455.8)
AS1	Subepidermal cells of cotyledon primordia	Byrne *et al.*, 2000	332.93 (±310)	43.28 (±40.6)
FIL/YAB1	Abaxial cotyledon primordia	Sawa *et al.*, 1999; Siegfried *et al.*, 1999	307.48 (±96.3)	52.67 (±71)

The mean signal value of the six replicate samples is shown for each tissue. The mean expression determined by LCM and GeneChip® analysis is shown. Standard Deviation is shown in parentheses.

microarray data to have a distinct spatial bias towards either the cotyledon or the root and then construct promoter::GUS fusions from these.

It was decided to target specifically putative transcription factors despite these often not having particularly high signal values. This approach will therefore provide a combination of microarray data validation and also uncover potential transcription factors important in regulating apical-basal polarity in the developing embryo.

All the data were collated together and a mean signal value for the cotyledon and root calculated for each gene. A ratio of mean cotyledon signal versus mean root signal was then calculated for each gene. This ratio, or 'fold-change', was then used to rank the genes based on whether they were highly expressed in the cotyledon relative to the root or the inverse. A sum of the Affymetrix detection call value (-1 absent, 0 marginal. 1 present) was also used to rank the data. A combination of 'fold change' and detection call was used to produce sub-sets of cotyledon- and root-biased genes. The selected gene lists were then annotated from a number of databases including MIPS (Mewes *et al.*, 1997; http://mips.gsf.de/), TAIR (Rhee *et al.*, 2003; http://www.arabidopsis.org/) and TIGR (http://www.tigr.org/db.shtml). A further level of sorting produced a list of putative transcription factors from which potential genes for promoter::GUS analysis were selected by hand.

Genes selected for promoter::GUS analysis based on LCM and GeneChip® analysis of torpedo stage are shown in Table 4.3 (Part 1). Three were relatively highly expressed in the cotyledon of torpedo embryos compared with the root pole, while three were relatively highly expressed in the root. The expression patterns in earlier stages of embryogenesis are also shown in the Table. In addition a number of promoter::GUS constructs were also created for genes selected from GeneChip® data obtained from LCM heart stage embryos by Casson *et al.* (2005), and these are shown in Table 4.3 (Part 2). All show relatively high expression levels in the heart-stage embryonic root.

Table 4.3 (Part I) Genes selected for promoter::GUS analysis (Torpedo stage).

Gene/construct designation	Description	Globular Apical	Globular Basal	Heart Cotyledon	Heart Root	Torpedo Cotyledon	Torpedo Root
At1g63900/A	putative RING zinc finger protein	193.4 (±130.8)	418.4* (±383.7)	525.2* (±765.9)	135 (±122.8)	1346.2 (±1152.1)	236.5* (±278.1)
At1g78160/B	putative RNA-binding protein	95.7 (±15.8)	139.2 (±64.9)	115.1 (±85.1)	59.5 (±15)	51.5 (±30.1)	216.2 (±204.1)
At5g14610/C	DRH1 DEAD box protein - like	331.9* (±531.6)	403.7* (±652.8)	2710.6 (±548.2)	395.7* (±405.8)	799.8 (±372)	259.4 (±244.8)
At5g43040/D	CHP-rich zinc finger protein-like	95.4 (±46.9)	102.9 (±44)	99.1 (±61.5)	48.4 (±9.1)	54* (±46.9)	135.3 (±119.4)
At5g50810/E	small zinc finger-like protein	682 (±116.8)	266.5 (±259.7)	165.1 (±188.3)	504.1 (±444.9)	798.3* (±857.5)	175.4 (±180.1)
At2g45050/F	putative GATA-type zinc finger transcription factor	28 (±25.7)	27.5 (±22)	38.9 (±46.3)	23.3 (±17.9)	6.6 (±5.5)	114.2 (±85)

The mean expression determined by LCM and GeneChip® analysis is shown. Standard Deviation is shown in parentheses. * Indicates an abnormally high mean signal value due to a large signal variation in one replicate.

Table 4.3 (Part II) Genes selected for promoter::GUS analysis (Heart stage)

Gene/construct designation	Description	Globular Apical	Globular Basal	Heart Cotyledon	Heart Root	Torpedo Cotyledon	Torpedo Root
At2g25420/Q	Putative WD-40 domain	10.2 (±2.8)	36 (±29.8)	85.8 (±117)	586.2 (±304.3)	58.7 (±48.2)	79.3 (±70.5)
At2g31510/R	Putative RING zinc finger protein	91.4 (±91.5)	60.6 (±49)	73.4 (±63.6)	834.6* (±809.7)	63.7 (±33.6)	65.7 (±46.7)
At3g60860/S	guanine nucleotide exchange factor - like protein	126.1 (±125.8)	151.1 (±175.65)	130.7 (±125)	2430.3 (±2675.4)	611.1* (±69.6)	148.9 (±166.2)
At5g45600/T	Putative YEATS domain /transcriptional activator	481.8 (±207.6)	413.3 (±495.6)	244.8 (±93.7)	2952.4 (±1987.6)	518.8* (±574.3)	455 (±481.5)

The mean expression determined by LCM and GeneChip® analysis is shown. Standard Deviation is shown in parentheses. * Indicates an abnormally high mean signal value due to a large signal variation in one replicate.

4.5.1 Creating promoter::GUS constructs

The generalised approach for creating promoter::GUS constructs is outlined in Figure 4.2. In brief, an approximately 2.5 kb genomic sequence upstream of the gene of interests ATG was cloned. This sequence was predicted to contain the full-length promoter driving gene expression in the wild-type. The promoter fragment was initially cloned into the TOPO vector (Figure 2.1) and sequenced to ensure the product was correct and that no errors had been accrued through PCR. The promoter fragment was then transferred to the binary vector pΔGUS-CIRCE, which contains the *β-glucuronidase* (GUS) reporter gene, for transformation into plants via *Agrobacterium tumefaciens*. This allows a comparison to be made between GUS expression driven by the promoter and the pattern predicted from LCM and GeneChip® analysis. The cloning of the promoter fragment into pΔGUS-CIRCE was accomplished using a slightly different method for construct T than that used for constructs A-S (See Appendix 3). All constructs were validated by sequencing.

4.5.1.1 PCR for promoter fragments

In order to clone the promoter from genomic DNA, two PCR reactions were carried out for constructs A-F. Two sets of primers were designed to amplify the promoter region. The first set of primers (designated GENE Forward Inner and GENE Reverse Inner), were designed to include a restriction enzyme site at each end of the amplified product to facilitate subsequent cloning. The second set of primers (designated GENE Forward Outer and GENE Reverse Outer), were designed outside the first set to amplify a region of genomic DNA including the 2.5 kb promoter fragment amplified by the first set. The outer promoter PCR used *Taq* polymerase (Section 2.10.1) and a reduced number of cycles to amplify the promoter region from a genomic DNA template. The inner promoter PCR used the proofreading Expand *Taq* polymerase (Section 2.10.2) and the product of the outer reaction as a template. This strategy was employed in an attempt to minimise PCR errors. For constructs Q-T only the inner reaction was used with the 2.5Kb promoter fragment amplified directly from a genomic DNA template using *Taq* polymerase.

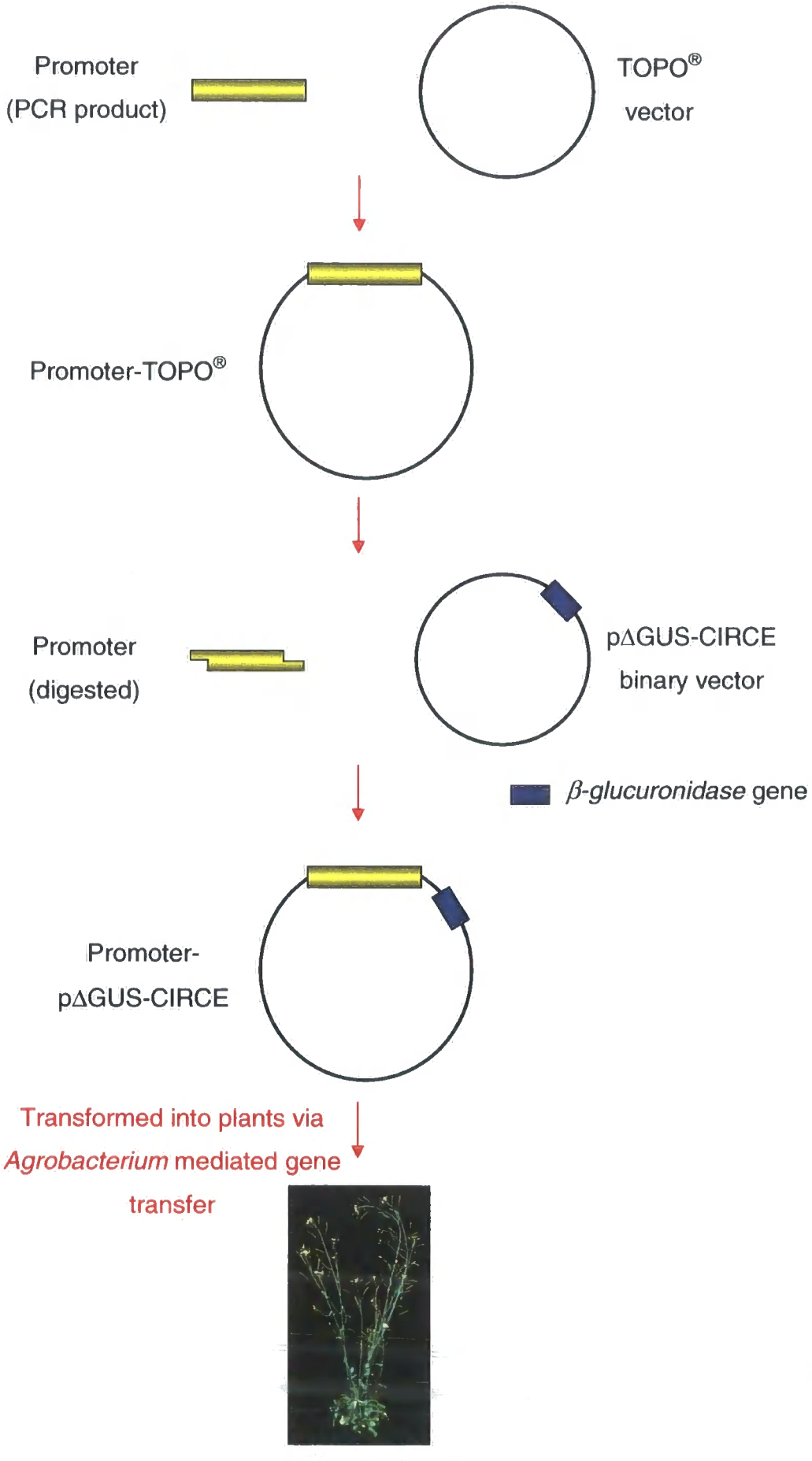
Figure 4.2 Approach taken to clone promoter::GUS constructs.

Constructs A-S:

- 2.5 kb promoter fragment PCR product with restriction enzyme (X) recognition sites at both ends.
- Promoter PCR product cloned into TOPO® vector.
- Promoter fragment digested out of TOPO® vector with relevant restriction enzyme (X).
- pΔGUS-CIRCE binary vector digested with relevant restriction enzyme (X) to provide 'sticky ends' for ligation.
- Promoter fragment ligated with pΔGUS-CIRCE binary vector.
- Construct transformed into plants via *Agrobacterium* mediated gene transfer.

Construct T:

- 2.5 kb promoter fragment PCR product cloned into TOPO® vector.
- Promoter fragment digested out of TOPO® vector using restriction enzyme recognition sites surrounding the TOPO® cloning site.
- 3' adenine residues added to both ends of the promoter fragment to provide 'sticky ends' for ligation.
- pΔGUS-CIRCE binary vector digested with Sma I and T-tailed to provide 'sticky ends' for ligation.
- Promoter fragment ligated with pΔGUS-CIRCE binary vector.
- Construct transformed into plants via *Agrobacterium* mediated gene transfer.


Promoter
(PCR product)
TOPO®
vector
Promoter-TOPO®
Promoter
(digested)
pΔGUS-CIRCE
binary vector
β-glucuronidase gene
Promoter-
pΔGUS-CIRCE
Transformed into plants via
Agrobacterium mediated gene
transfer

1) Outer promoter PCR

The Outer PCR was carried out as described in Section 2.10.1, using GENE Outer primers, genomic DNA template and *Taq* polymerase. An aliquot of the reaction mix was then run out on a gel as described in Section 2.9, to confirm the product was of the expected size. The product was then purified using the High Pure PCR cleanup kit (Roche) as described in Section 2.8.5.2 and quantified by running out an aliquot of the purified product on a gel with DNA markers.

2) Inner promoter PCR

The Inner PCR was carried out as described in Section 2.10.2, using GENE Inner primers, the Outer PCR reaction product as template and Expand *Taq* polymerase (Roche). An aliquot of the reaction mix was then run out on a gel as described in Section 2.9, to confirm the product was of the expected size. The product was then purified using the High Pure PCR cleanup kit (Roche) as described in Section 2.8.5.2 and quantified by running out an aliquot of the purified product on a gel with DNA markers.

4.5.1.2 Addition of 3' Adenine Overhangs to PCR Products

PCR products generated by DNA polymerases that lack the ability to add template independent 3' adenine residues (such as *Tgo* the proofreading DNA polymerase used in the Expand high fidelity PCR system) require the addition of these bases before they could be used in TA overhang-dependant cloning reactions, such as cloning into the TOPO vector (see Section 2.11.7.1). PCR products for constructs A-F required the addition of 3' adenine overhangs.

4.5.1.3 Cloning the promoter into the TOPO vector

Promoter fragments were cloned into the plasmid vector pCR®2.1-TOPO® (see Figure 2.1) using the one-step "TOPO® cloning" strategy (as described in Section 2.11.7). This provided an intermediate stopping point in the overall cloning strategy, allowing the promoter fragment to be sequenced to ensure errors had not been accrued by PCR, and to ensure restriction sites were present for subsequent cloning.

4.5.1.4 Cloning the promoter into the Binary vector pΔGUS CIRCE

pΔGUS-CIRCE is a pBIN19 derivative plasmid containing the *β-glucuronidase* gene downstream of a multiple cloning site as shown in Figure 2.2 (Jefferson, 1987). In order to introduce the promoter fragment into the vector, restriction digestion by endonucleases was performed using the following enzymes on each promoter insert and vector for ligation.

1. Constructs A-F and R were digested with *Bam*H I as described in Section 2.11.1, and run out on a gel.

2. Constructs Q and S were digested with *Sal* I as described in Section 2.11.1, and run out on a gel.

Digests were run out on a gel to ensure complete restriction had occurred. The promoter inserts were ligated into pΔGUS-CIRCE (digested with the respective restriction enzyme), transformed into TOP10 (constructs A, C, E, R, S,) or DH5α (constructs B, D) cells and positives selected. This colony PCR was conducted using a nested primer for the *β-glucuronidase* gene and either the forward inner primer (constructs R-T) or a forward primer designed inside the promoter sequence (designated Forward Internal 3), which resulted in a shorter product (~1 kb) (constructs A-E). A single positive colony was selected from each transformation and its DNA sequenced to ensure there were no errors prior to transformation into plants.

No colonies containing successfully ligated promoter-vector constructs were obtained for constructs F and Q.

The cloning strategy employed for construct T utilised a TA overhang-dependant cloning reaction. The T promoter fragment was first cloned into the TOPO vector to confirm its identity and assess for sequencing errors. It was digested out of the TOPO vector using *Bam*H I and *Xba* I, and 3' adenine residues were then added, thus providing a 'sticky end' for ligation. The pΔGUS-CIRCE vector was digested with *Sma* I and T-tailed as described in Section 2.11.4, thus creating a single, overhanging 3' deoxythymidine (T) residue. This allowed the promoter PCR insert to ligate efficiently with the pΔGUS-CIRCE. The ligated promoter-vector construct was transformed into TOP10 cells and positives selected. A single positive colony was selected from the transformation and its DNA sequenced to ensure there were no errors prior to transformation into plants.

4.5.1.5 Introducing the promoter-GUS construct into plants

Plasmid DNA was isolated and purified from positive *E.coli* colonies. Constructs were introduced into *Agrobacterium* by electroporation, as described in Section 2.13.1. Plants were transformed using the dipping method, as described in Section 2.13.2. Transformed plants were then grown to maturity and their seed collected and germinated on 35 mg/l kanamycin to select for transformants.

Kanamycin selection for transformants generated resistant plants for the promoter constructs B, C, E, R and T. No plants' containing the promoter constructs for A, D and S were successfully obtained. Resistant lines were removed from Petri dishes and planted in soil to set seed.

4.5.2 Analysis of GUS expression in transgenic plants

In order for meaningful comparisons to be made between different stages of development post germination a standardised GUS histochemical staining procedure was used, as described in Section 2.4. In order to ascertain the optimal length of time for staining in each promoter::GUS construct, plants were stained over a time-course (see Section 2.4.2). Optimal staining for all GUS line seedlings was found to be between 4 and 8 hours, and overnight staining for embryos.

No formal segregation analysis was undertaken, however, the GUS staining patterns described were observed at a high frequency in multiple independent transformed lines for each construct (with the exception of construct B).

4.5.2.1 Analysis of GUS activity in mature tissue

Seed from resistant lines was germinated on 35 mg/l kanamycin and seedlings removed from Petri dishes at 3, 5, 7 and 12 days post germination for histological analysis of GUS expression. Penetration of the seedling by the histochemical buffer and substrate was achieved by subjecting seedlings to a 20 minute vacuum infiltration prior to incubation at 37°C for the optimal period of time.

A resistant line for construct B showed staining in the anther, stigma and silique, no analysis of GUS staining in the seedling was carried out for this line (Figure 4.3).

Resistant lines for construct C showed GUS staining exclusively in the stomatal guard cells (Figure 4.4).

Resistant lines for construct E showed faint GUS staining in the vascular associated cells of the cotyledon and true leaf and also stronger staining in the stipules of the 7dpg seedling. Strong GUS staining was observed in the primary root and lateral root primordia (Figure 4.5).

Resistant lines for construct R showed constitutive GUS staining throughout the vasculature of the seedling cotyledons and root (Figure 4.6)

Resistant lines for construct T showed a changing pattern of GUS staining in the seedling with strong staining observed in the root at 3 dpg becoming constitutive in the root and arial parts by 5 dpg. At 7dpg diffuse GUS staining is observed throughout the arial parts and strong staining is observed in the root and lateral root primordia. No staining is observed in the hypocotyl (Figure 4.7).

4.5.2.2 Analysis of GUS activity in the embryo

Following growth in soil, siliques of various developmental stages were removed from plants and analysed for GUS expression by histology. In order to stain embryos, siliques were dissected open using a fine-pointed needle, and each seed coat pierced individually. Penetration of the seed by the histochemical buffer and substrate was achieved by subjecting dissected siliques to a 20 minute vacuum infiltration prior to incubation overnight at 37°C.

A single heart-stage embryo showing GUS staining was obtained from a resistant line for construct B, the staining appeared to be present in both the cotyledon and root domains (as well as the endosperm) (Figure 4.3). Further characterisation is required to confirm this pattern for construct B.

Resistant lines for construct C showed GUS staining in the cotyledons and hypocotyl of the cotyledonary stage embryo, with occasional diffuse staining observed in the root (Figure 4.4).

Resistant lines for construct E showed a changing pattern of GUS staining with constitutive staining observed at the heart-stage of embryogenesis, expression is then lost in the root as the embryo enters the torpedo-stage of embryogenesis (Figure 4.5).

Figure 4.3 Promoter::GUS analysis of At1g78160, a putative RNA-binding protein (construct B).

A Anther, showing GUS staining in the pollen grains and tapetum

B Stigma, showing GUS staining in the papillae

C Silique, showing GUS staining at the junction of the placenta and funicle (silique)

D Heart-stage embryo and endosperm

E Silique, showing GUS staining of the vascular strands of the receptacle and placenta

Scale bars in A-D represent 50 μm, scale bar in E represents 500 μm.

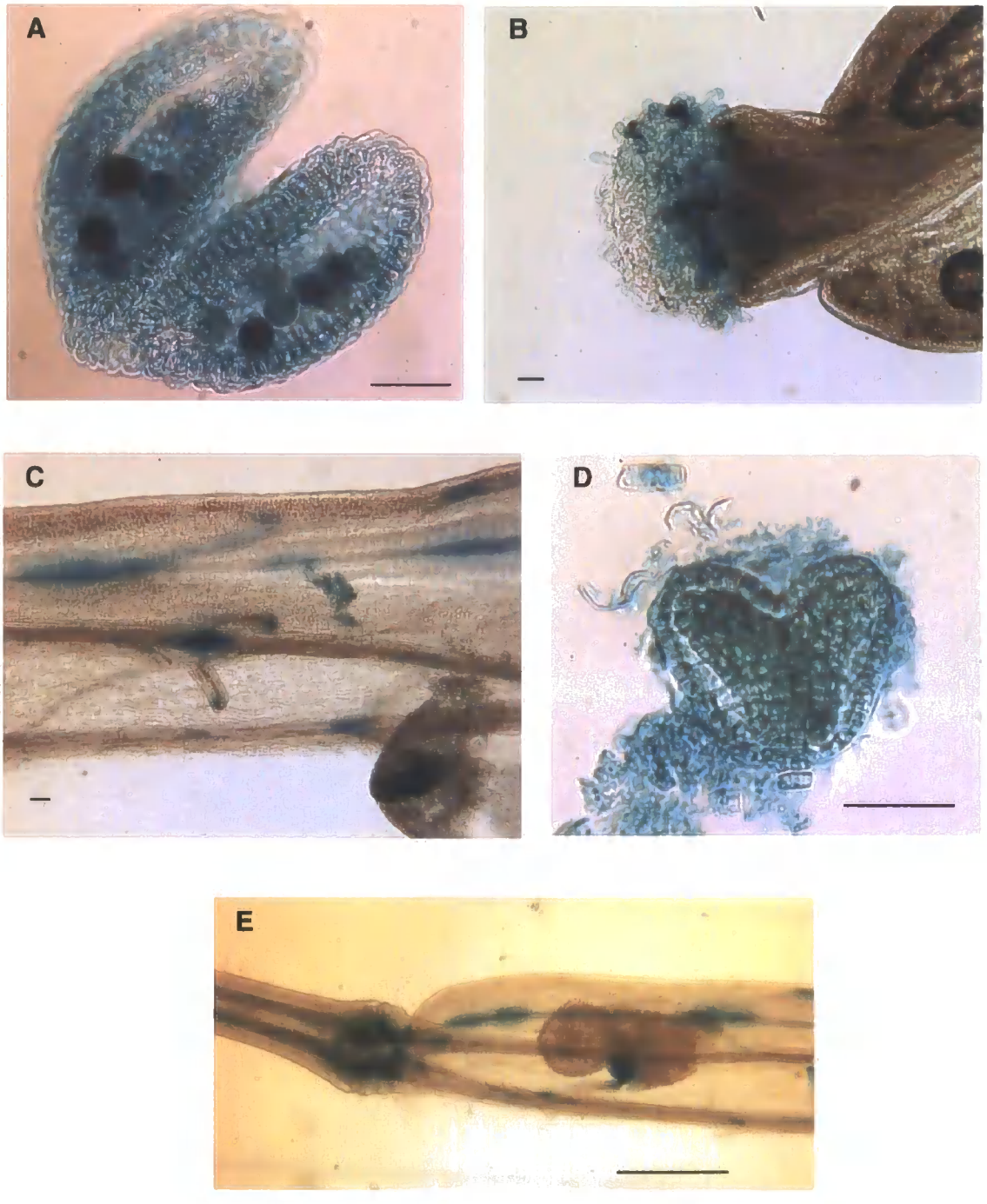

A
B
C
D
E

Figure 4.4 Promoter::GUS analysis of At5g14610, a DEAD box RNA helicase family-like protein (construct C).

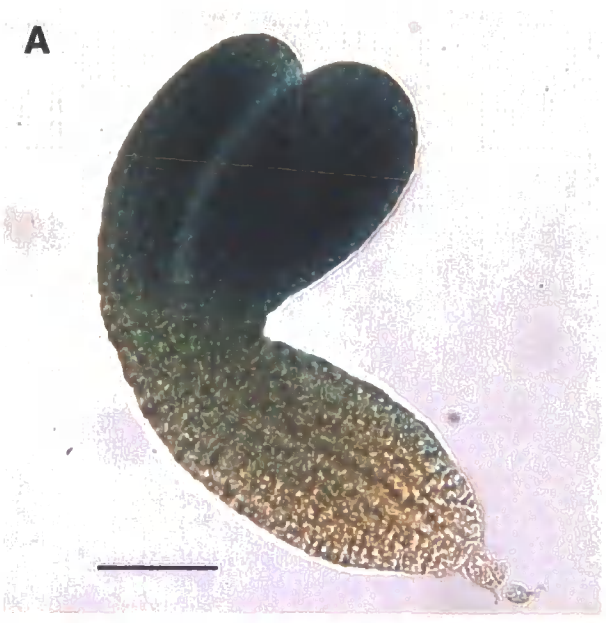
A Cotyledonary-stage embryo

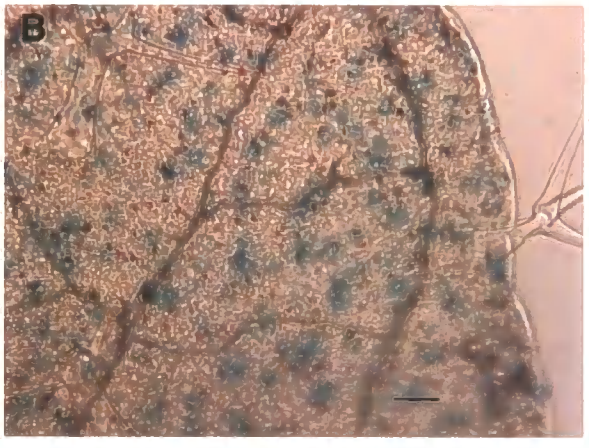
B 7dpg seedling, true leaves showing GUS staining in the stomatal guard cells

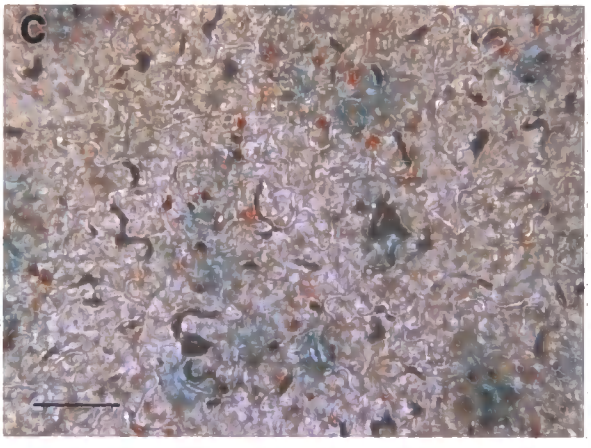
C 7dpg seedling, cotyledon adaxial surface showing GUS staining in the stomatal guard cells

D Silique, showing GUS staining in the stomatal guard cells

All scale bars represent 50 μm.


A


B


C

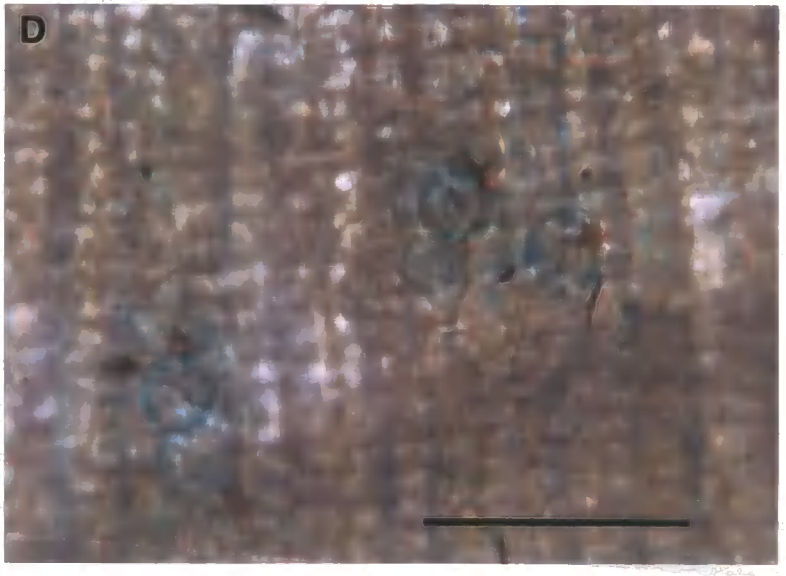

D

Figure 4.5 Promoter::GUS analysis of At5g50810, a small zinc finger-like protein (construct E).

A Late heart-/early torpedo-stage embryo

B Torpedo-stage embryo

C 7 dpg seedling, showing GUS staining in vascular-associated cells around the primary mid-vein of the cotyledon

D Silique, showing GUS staining in the funicles and ovules

E 7 dpg seedling, showing GUS staining in the stipules

F 7dpg seedling, showing GUS staining in the primary root

G 7dpg seedling, showing GUS staining in a lateral root primordium and the stele of the mature root

All scale bars represent 50 μm.

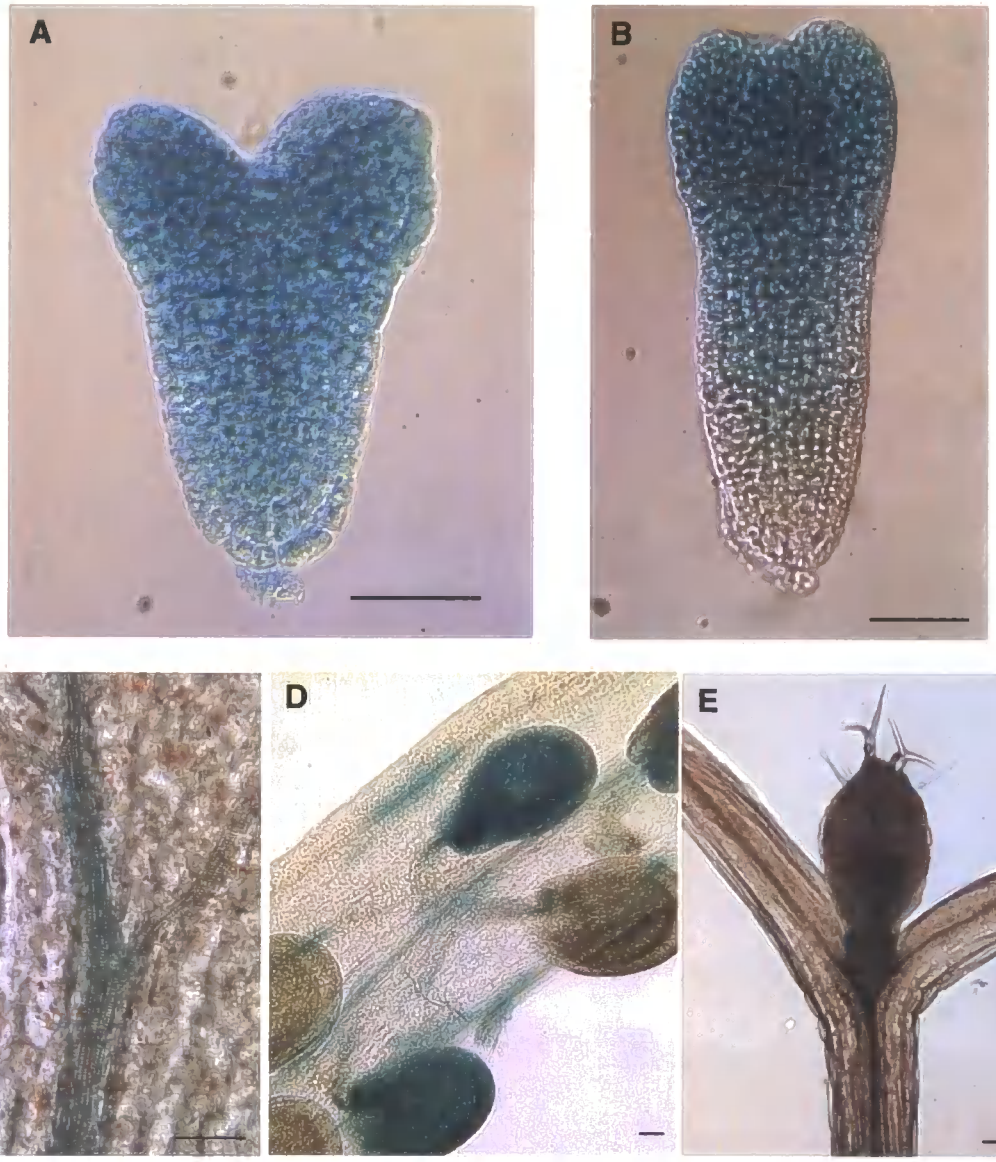

A
B
C
D
E
F
G

Figure 4.6 Promoter::GUS analysis of At2g31510, a putative RING zinc finger protein (construct R).

A Cotyledonary-stage embryo (strong expression in root pro-vascular strand)

B Silique, showing GUS staining in the recepticle, placenta and funicles

C 7dpg seedling, showing GUS staining in the primary mid-vein of the true leaves

D 7dpg seedling, showing diffuse GUS staining where the vascular strands dissociate and in the cotyledon stomatal guard cells

E 7dpg seedling, showing GUS staining in the root tip (pro-vascular tissue)

F 7dpg seedling, showing GUS staining in a lateral root primordium (pro-vascular strand)

Scale bars in A, C-F represent 50 μm; scale bar in B represents 500 μm.

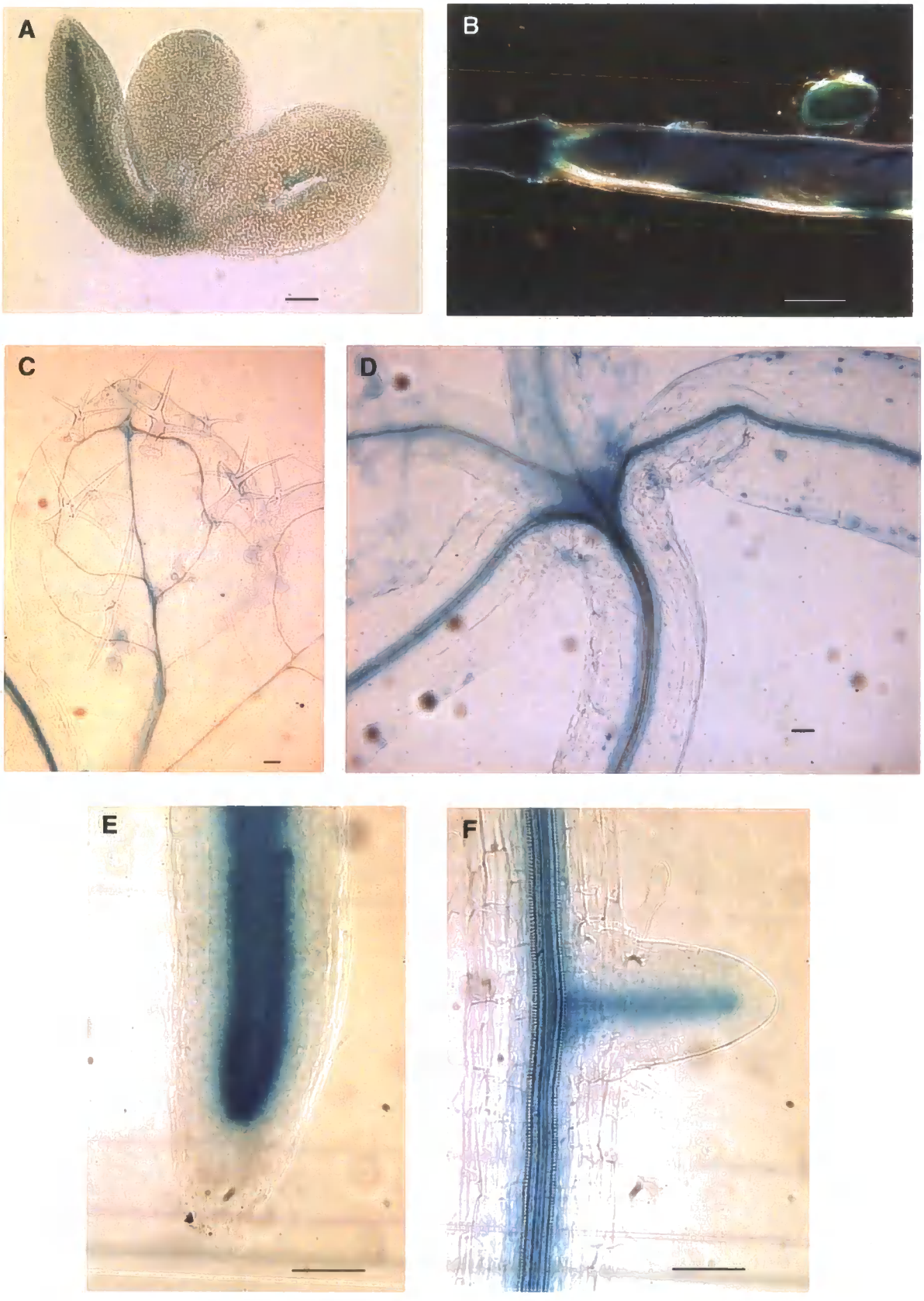

A
B
C
D
E
F

Figure 4.7 Promoter::GUS analysis of At5g45600, a putative YEATS domain/transcriptional activator protein (construct T).

A Cotyledonary-stage embryo

B Silique, showing diffuse GUS staining

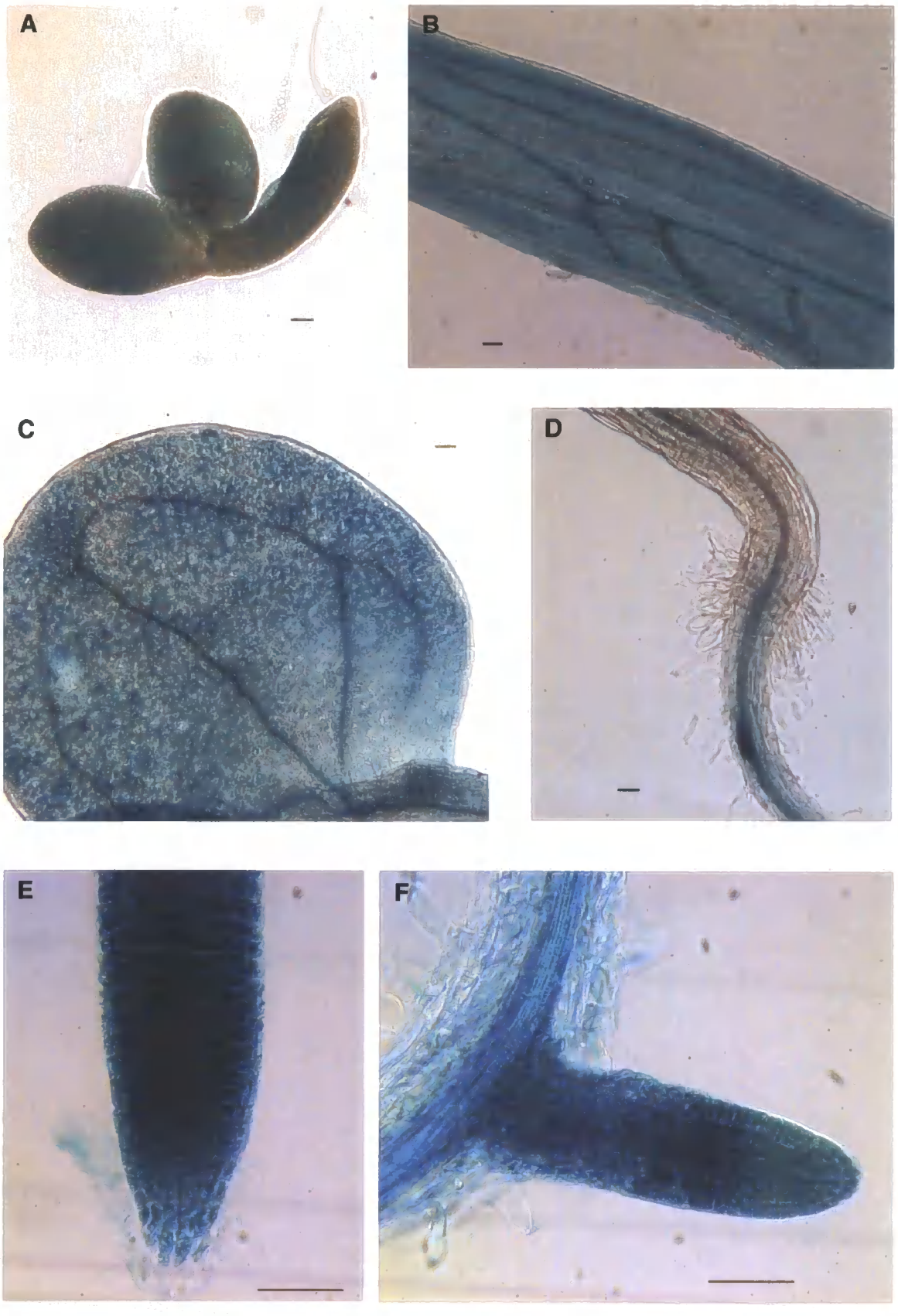
C 7dpg seedling, showing diffuse GUS staining in the true leaf, particularly vascular-associated tissue

D 7dpg seedling, no staining in hypocotyl, GUS staining of the hypocotyl/root junction and root

E 7dpg seedling, showing GUS staining in the primary root tip

F 7dpg seedling, showing GUS staining in the primary root stele, and throughout a lateral root primordium

All scale bars represent 50 μm.


A
B
C
D
E
F

Cotyledonary-stage embryos of resistant lines for construct R showed a constitutive pattern of GUS expression with more intense staining observed in the root pro-vascular strand (Figure 4.6).

Cotyledonary-stage embryos of resistant lines for construct T showed a constitutive pattern of GUS expression throughout (Figure 4.7).

4.6 Functional analysis of differentially regulated putative transcription factors

The creation of promoter::GUS constructs for differentially regulated transcription factors provided a means of validating the microarray data produced from LCM and GeneChip® analysis, while simultaneously suggesting genes of putative importance in the regulation of apical-basal polarity in the plant embryo. To complement this work on expression pattern, it was decided to investigate the functional role of the putative transcription factors selected. One means of investigating gene function *in vivo* is through the elimination or disruption of gene expression. Such disruption can be achieved through the insertion of an *Agrobacterium tumefaciens*-derived T-DNA fragment into the genome within the gene of interest, therefore inhibiting its transcription and consequently the production of a functional protein. As previously described in Section 1.1.1.4, the naturally occurring system of T-DNA insertion has been optimised so that tumours are not induced and therefore resulting phenotypes can be attributed to the disruption caused by integration into the genome.

T-DNA integration is largely random in nature; therefore the T-DNA flanking region must be amplified by PCR using a T-DNA border specific primer. Alignment of the T-DNA flanking region to the genomic sequence is then used to identify the location of insertion. A number of significant T-DNA insertion line collections have been established and made available to the research community (http://signal.salk.edu/cgi-bin/tdnaexpress).

Analysis of lines disrupted in the genes of interest was hoped to provide an insight into gene function through the analysis of any developmental effects observed.

4.6.1 Identification of knockout lines

It was decided to investigate the potential function of a range of putative transcriptional factors selected from the microarray data as showing differential expression. A search of the SALK line collection (http://signal.salk.edu/cgi-bin/tdnaexpress; Alonso *et al.*, 2003) revealed potential T-DNA insertion lines in the putative transcription factors of interest as shown in Table 4.4. Insertions in exons were predicted to result in either a null mutation or a truncated transcript. Insertions in other regions, such as the promoter were predicted to result in either a null mutation (through disruption of essential promoter function) or reduced expression (Krysan *et al.*, 1999).

SALK line seed was processed as described in Section 2.3 and cultivated prior to DNA extraction as described in Section 2.8.3.2. Genotyping was accomplished using the method described in Section 2.10.4 and shown in Figure 2.3. Two PCR reactions were used to genotype each line. The primers in the first reaction were designed around the T-DNA insertion site; production of a PCR product indicated the presence of an uninterrupted wild-type allele, as the product size if a T-DNA insert were present was prohibitive. The second reaction contained a primer specific to the left border of the T-DNA and therefore production of a product indicated that an insert was present. Plants were identified that were homozygous for the insertion and therefore lacking a functional copy of the gene of interest. Sequencing of the PCR product obtained from the insert-specific reaction was carried out for each SALK line to confirm the location of insertion.

Table 4.4 SALK T-DNA insertion lines for functional analysis of putative transcription factors.

To investigate the potential function of a range of putative transcriptional factors, an analysis of T-DNA insertion lines obtained from the SALK line collection (Alonso *et al.*, 2003) was carried out. Insertions in exons were predicted to result in either a null mutation or a truncated transcript (precise insert location indicated in parentheses, e.g. insertion in the 4th exon, of 11 total exons). Insertions in the promoter were predicted to result in either a null mutation (through disruption of essential promoter function) or reduced expression (Krysan *et al.*, 1999).

Gene	SALK Line	Predicted Insert Location	Location Confirmed by Sequencing
At1g63900	SALK_063571	Exon (4 of 11)	✓
At1g78160	SALK_135981	Promoter	✓
At5g14610	SALK_116644	Exon (3 of 11)	✓
	SALK_068359	Exon (7 of 11)	✓
At5g43040	No SALK Line available		
At5g50810	SALK_144520	Exon (1 of 2)	✓
	SALK_012862	Exon (2 of 2)	✓
At2g45050	SALK_508845	Exon	WT Only
At2g25420	SALK_068825	Exon (5 of 15)	✓
	SALK_131239	Exon (6 of 15)	✓
At2g31510	SALK_132004	Exon (9 of 15)	✓
At3g60860	SALK_033446	Exon (9 of 11)	✓
At5g45600	SALK_106430	Exon (4 of 11)	✓

4.6.2 Effects of T-DNA insertion on plant morphology

Initial screening of plants homozygous for SALK insertions did not reveal any obvious differences from wild-type plants grown under identical conditions. This observation is based upon a comparison of general morphology including factors such as growth rate, leaf number/shape, and time of bolting being considered. An analysis of general embryo morphology was also carried out. For lines with a predicted root phenotype vertical plate measurements of root length were conducted (as described in Section 2.3.1), no significant difference was observed between the homozygous SALK T-DNA insertion population and the wild-type population.

4.7 Expression patterns of known *EMB* genes

Genes which under normal conditions are required for viability, and when disrupted cannot be passed on to subsequent generations can be considered essential. The precise number of such genes with essential functions during embryogenesis has not yet been established, however it is now estimated that 500 to 1000 genes in *Arabidopsis* produce an embryo-defective phenotype when mutated (Franzmann *et al.*, 1995; McElver *et al.*, 2001). Tzafrir *et al.* (2004) have described a current collection of 220 *EMB* genes required for normal embryo development, potentially representing between 25 and 50% of the eventual total. The embryonic expression pattern for many of the genes in this collection have not been characterised, it was therefore decided to analyse the GeneChip® data for these essential genes in an attempt to add a spatial dimension to our knowledge.

Applying the two arbitrary signal cut-off values as before, it was determined that approximately 84% of the *EMB* genes were expressed in at least one of the three tissue types at a signal threshold of 75, rising to almost 96% at a signal threshold of 40. An analysis into the spatial expression of these genes revealed that around 76% were present in all tissue types at the lower signal threshold, as is shown in Figure 4.8. The 4% of *EMB* genes deemed not to be expressed at the lower signal threshold could potentially be expressed in the hypocotyl

Figure 4.8 Detection of EMB gene expression in torpedo-stage embryos.

The Venn diagram indicates the number of EMB genes (Tzafrir *et al.*, 2004) expressed (mean signal value >40) in the cotyledons, root and shoot apical meristem. The overlapping regions indicate the number of genes expressed in two tissue types or all tissues. The numbers in parentheses indicate the number of genes not detectably expressed in a particular tissue type. The number in the box indicates the number of genes not expressed in any tissue type at a mean signal cut-off value of 40.

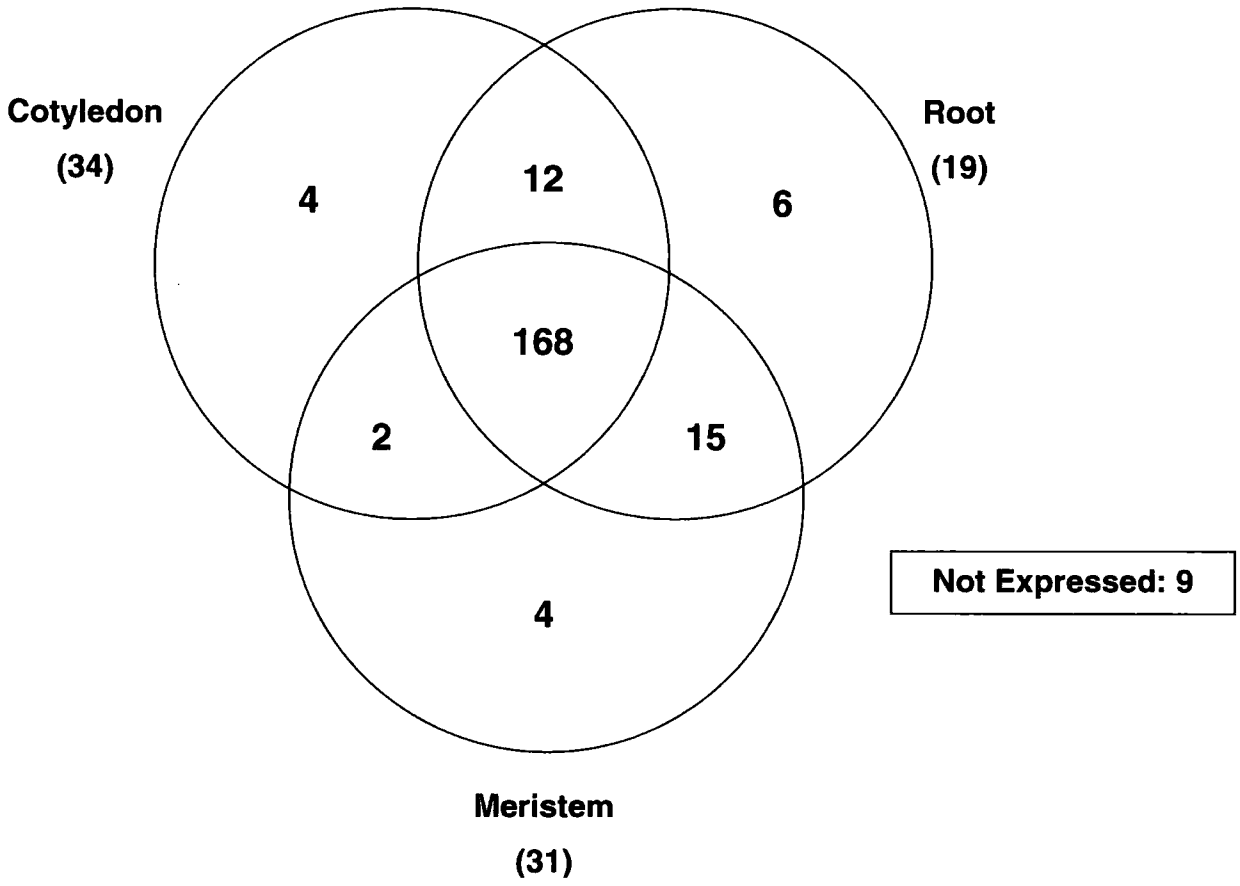

Cotyledon
(34)
Root
(19)
4
12
6
168
2
15
4
Meristem
(31)
Not Expressed: 9

Table 4.5 *EMB* genes determined to show spatial differential expression during the torpedo-stage of embryogenesis.

Genes selected from expressed *EMB* genes (mean signal value >40) that show at least a fourfold difference in expression between embryo zones, filtered for outlier and inconsistent signal values.

Cot – Cotyledon
Meri - Shoot Apical Meristem
Root - Root

Gene	Gene symbol	Fold difference
At2g34650	PID	15x Cot > Root
At5g18580	FS1	10.1x Cot > Root 9.1x Meri > Root
At1g80070	SUS2	5.6x Cot > Root 6.1x Meri > Root
At4g30090	EMB 1353	11.5x Root > Cot 17.8x Meri > Cot
At3g20070	TTN9	6.6x Root > Cot 20.2x Meri > Cot
At3g20630	TTN6	4.1x Root > Cot
At1g62360	STM	62.2x Meri > Cot 19.3x Meri > Root
At1g78580	TPS1	6.1x Meri > Cot
At2g21710	EMB 2219	5.6x Meri > Cot 4.7x Meri > Root
At1g80410	EMB 2753	8x Meri > Cot 5.1x Meri > Root
At2g31340	EMB 1381	4.5x Meri > Cot
At5g07280	EXS	4.5x Meri > Root

region, which was not sampled, or at another stage of embryogenesis. Indeed a comparison with GeneChip® data presented by Casson *et al.* (2005) showed that two of the non-expressed genes were present at earlier stages of embryogenesis, At4g21130 at the globular stage and At2g45690 at both the globular and heart stages, albeit at threshold levels.

From analysis of the data it was possible to identify *EMB* genes with differential expression patterns between tissue types (see Table 4.5). Significant examples include, At2g34650 (*PID*), which is approximately 15-fold more abundant in the cotyledons than the root, and At1g62360 (*STM*), which is approximately 62-fold more abundant in the shoot apical meristem than in cotyledonary tissue.

4.8 Discussion

The previous chapter demonstrated that LCM could successfully be applied to tissue of a torpedo-stage embryo, and that in combination with an amplification protocol produce high quality aRNA. This was confirmed by RT-PCR analysis for genes with known spatially distinct expression patterns, verified by *in situ* hybridisation.

Sufficient aRNA was generated to undergo microarray analysis to allow a global analysis of gene expression levels. In order to allow a measure of confidence to be attributed to the data generated it was deemed of critical importance to validate using a number of different methods. In addition to comparing the expression patterns exhibited by a number of individual genes, the entire data set was compared to a number of current studies to ascertain whether the total number of genes estimated to be expressed was comparable.

4.8.1 Estimation of the number of genes expressed

The imposition of an Affymetrix MAS 5.0 signal cut-off value has been employed in two recent studies as a means to providing an estimation of the number of genes expressed in a tissue of interest (Birnbaum *et al.*, 2003; Casson *et al.*, 2005).

Birnbaum *et al.* (2003) conducted a study of gene expression in the *Arabidopsis* root using a rapid protoplasting technique (see Section 1.4.2). Concerns have been raised as to the use of this technique in gene expression profiling, due to changes in gene expression induced as a consequence of the manipulations involved in protoplasting (Grosset *et al.*, 1990). Birnbaum *et al.* (2003) compared the expression profiles in protoplasted roots with untreated roots and concluded that the technique did not change global gene expression profiles to any significant degree. However, this study did uncover several hundred genes, which were consistently induced by the protoplasting treatment, which they then removed from their analysis. Using a number of genes with previously documented expression patterns, Birnbaum *et al.* (2003) calculated a signal cut-off value that represented a minimum value to confer presence over the level of background noise, and this was set at 75. Using this cut-off they estimated that 10,492 genes were expressed in the root, and this corresponds to ~46% of the genes represented on the ATH1 GeneChip. Casson *et al.* (2005) also applied this signal cut-off value to their GeneChip data obtained from the apical and basal regions of globular- and heart-stage embryos, finding between 8027 and 10,591 genes (36-47%) to be expressed. Using the cut-off value of 75 for the mean value of the replicates analysed, my data indicate that between 8353 and 11,690 genes (~37-51%) are expressed in the three tissue types of the torpedo stage embryo analysed. My data are in a very comparable range to these previous studies and partially bridges the temporal gap between them, demonstrating that at this arbitrary cut-off value a similar number of expressed genes are estimated throughout embryogenesis and in the mature *Arabidopsis* root.

In a similar way to that of Birnbaum *et al.* (2003), Casson *et al.* (2005) used the signal values of a number of genes with previously documented expression

patterns in embryogenesis to calculate a signal cut-off value that represented a minimum value to confer presence over the level of background noise. The signal cut-off value proposed was 40, with the argument for this lower value being the low mean signal values obtained for *AINTEGUMENTA* in the apical region of the globular-stage embryo and *MONOPTEROS* in the root pole of the heart-stage embryo, both genes with documented expression patterns in these regions confirmed by *in situ* hybridisation (Elliot *et al.*, 1996; Hardtke and Berleth, 1998). Using the signal cut-off value of 40, Casson *et al.* (2005) predicted that up to 65% of genes are expressed, this compares with a predicted range of expressed genes of 59-77% in the tissue regions analysed in the torpedo-stage embryo.

The estimation of gene expression threshold values is a useful tool to compare tissue regions superficially; however its potentially arbitrary nature is highlighted by a study of laser microdissected slender embryonic structures from a mouse. In a study designed to assess differences between T7-amplified and non-amplified samples using two-colour cDNA microarrays, a high correlation was found for high-abundance genes, but the correlation decreased markedly for low-abundance genes particularly at levels approaching background. This decrease in correlation was not between the amplified and non-amplified samples but across all experiments suggesting variability in hybridisation to the microarray in the low-abundance region (Scheidl *et al.*, 2002). Despite the difference in microarray technology used between the Scheidl *et al.* (2002) study and that used by Birnbaum *et al.* (2003) and Casson *et al.* (2005), the fundamental point remains valid, that in the very low-abundance region a greater degree of relative variability is potentially possible than at high-abundances. Therefore while this low-abundance region contains a wealth of potentially expressed genes, as shown by the increase in number of predicted expressed genes between the two signal cut-off values, it must be assumed that it is also affected by variability in hybridisation and therefore a specific yes-no cut-off can give only an estimation of expression threshold.

4.8.2 Comparison with published expression patterns of embryonic genes

A degree of variation in the signal value is found between replicate GeneChip data, and this could be attributed to either biological variation between samples, or to technical variability.

In order to ensure a normal spread of biological variation in my data, each replicate sample was obtained from an independent group of plants grown over sequential periods of time and at different positions in the growth room. Some of the variation observed in the signal value could be directly attributed to this strategy adopted to provide a representative spread of normal biological variation.

King *et al.* (2005) performed a study into variability of gene expression detectable using oligonucleotide arrays, with a view to assessing the reproducibility of RNA amplification. They observed a degree of biological variation between non-diseased human breast control samples but also suggested some technical variability might be resulting from decreased amplification among low abundance transcripts. In contrast, Scheidl *et al.* (2002) suggest that the technical variability in the low abundance region results from the microarray system and not from the amplification. Either way, as the variations observed in my data are most prevalent in genes with low-abundance expression values (i.e. those around the cut-off thresholds imposed), much of this variation can be attributed to the sensitivity of the system in this region, highlighting the requirement of replicates to be able to derive more useful information from low-abundance expression values.

Given the presence of variation in the data, whether biological or technical, it is of critical importance to provide a framework of validation so as to have confidence in the microarray data. One such approach to validation is to compare the expression data obtained from the amplified RNA of laser microdissected torpedo-stage embryos with expression pattern data previously reported for embryonic genes in the literature. As the expression patterns of genes in the literature have been confirmed by such techniques as *in situ*

hybridisation and promoter::GUS analysis then this would provide a very powerful form of validation if they were in agreement with the microarray data.

RT-PCR for genes of known expression pattern was performed on 3rd round aRNA as a means to validating the laser capture microdissection and amplification protocols (see Section 3.5). The results of this analysis correlated very closely with those expected based on the published expression patterns of the genes investigated, confirming the success of the methods used. Therefore, an analysis of the GeneChip® expression data for these genes would provide clear validation for the microarraying step alone.

Two genes were investigated whose expression domains encompassed the cotyledons but not the root, namely *AINTEGUMENTA (ANT)* and *PINOID (PID). aintegumenta (ant)* was isolated from *Arabidopsis* as a female-sterile mutant, defective in ovule and floral development (Elliott *et al.*, 1996). The wild-type *ANT* gene was cloned and shown to be a member of the *APETALA2*-like family of transcription factors (Elliott *et al.*, 1996; Klucher *et al.*, 1996). *ANT* expression during embryogenesis was examined by *in situ* hybridisation, and by the torpedo stage expression was found to be restricted to an internal region of the cotyledons with no expression in either the hypocotyl or the root (Elliott *et al.*, 1996). The GeneChip® data (Table 4.2) show *ANT* expression to be over 30-fold higher in the cotyledons (mean signal value, 647.28) compared to the roots (mean signal value, 20.73). Using a signal cut-off value of 40 to confirm the presence or absence of transcripts, *ANT* is expressed in the cotyledons but not in the roots, which is in agreement with the *in situ* hybridisation analysis and the RT-PCR performed on 3rd round aRNA.

pinoid (pid) displays an aberrant shoot phenotype with defects in the cotyledons and floral organs resembling that of the *pin-formed (pin)* mutant that is disrupted in polar auxin transport (Christensen *et al.*, 2000; Benjamins *et al.*, 2001). *PID* has been cloned and encodes a member of a serine-threonine protein kinase family (Christensen *et al.*, 2000). PID induces the apical-basal targeting of PIN1 and thus acts as a mediator of auxin flow, creating local gradients required for patterning (Friml *et al.*, 2004). *PID* was shown by *in situ* hybridisation to be transiently expressed during embryogenesis, with expression in the cotyledons

detectable until the mid-torpedo-stage, and expression absent in the root (Christensen *et al.*, 2000). This expression pattern is again in agreement with my GeneChip® data. It is of interest to note that the mean signal value in the root would be classed as present using the signal cut-off values suggested by both Birnbaum *et al.* (2003) and Casson *et al.* (2005), yet has been proposed as being absent in that tissue by *in situ* hybridisation. This highlights the problems associated with imposing a cut-off value due to variability in hybridisation at low-abundance levels and the requirement to validate findings using established methods.

AtPIN4, a member of the PIN family of auxin efflux facilitators is expressed in the root meristem throughout embryogenesis and in the mature plant, and was selected as the root-specific gene marker for RT-PCR analysis of 3rd round aRNA (Friml *et al.* 2002). The GeneChip® data once again demonstrate the difficulties presented by technical variability at low-abundance levels. There was observed a definite increase in *AtPIN4* expression in the root compared to the cotyledon, but not a particularly large differential expression was found, and a relatively high signal value was detected for the cotyledon region. Whole mount *in situ* immunolocalisation analysis of AtPIN4 occasionally resulted in weak staining in the apical embryonic region due to cross-reaction with other AtPIN proteins. The structural similarity between members of the gene family might also result in the background hybridisation variability observed on GeneChip® cotyledon signal value (Friml, 2002).

The *PIN* family represents a good example of genes with differential spatial and temporal expression patterns during embryogenesis. At the 16-cell stage *AtPIN1* is expressed in all cells of the embryo-proper in a non-polar fashion, with expression becoming restricted to the basal end of vascular precursor cells by the torpedo-stage (Steinmann *et al.*, 1999; Friml *et al.*, 2003); this constitutive expression pattern is reflected in my GeneChip® data. *AtPIN3* mRNA expression is first detected by promoter::GUS and *in situ* hybridisation in the root pole of the heart-stage embryo, suggesting that unlike the other three PIN family members expressed in the embryo (*PIN1, PIN4* and *PIN7*) it does not have a role in early apical-basal pattern formation (Friml *et al.*, 2003; Jenik

et al., 2005). The basal expression pattern is represented in my GeneChip® data with a greater differential expression to that exhibited by *AtPIN4.*

In addition to the *ACT3* constitutive control, *BODENLOS/IAA12 (BDL)* and *MONOPTEROS (MP)* were selected as constitutively expressed genes for RT-PCR analysis of the 3rd round aRNA. The *bdl* mutant was isolated from a screen of seedling phenotypes resulting from EMS mutagenesis (Hamann *et al.*, 1999). Development of the primary root meristem is impaired in *bdl* mutants, with the deviation from normal development observed from the two-cell stage of embryogenesis onwards (Hamann *et al.*, 1999). mRNA *in situ* expression analysis of the *BDL* transcript show it to be restricted to the vascular precursor cells by the torpedo-stage of embryogenesis, a constitutive expression pattern confirmed by my GeneChip® data. *MP* encodes *ARF5*, an auxin-responsive transcription factor shown to interact with *BDL*, with co-expression throughout early and mid-embryogenesis (Hardtke and Berleth, 1998; Ulmasov *et al.*, 1999; Hamann *et al.*, 2002). My GeneChip® data are in accordance with the constitutive expression pattern in vascular precursor cells confirmed by *in situ* hybridisation (Hamann *et al.*, 2002).

A number of known embryonic genes with expression patterns confirmed by *in situ* hybridisation and promoter::GUS studies were analysed. In all cases my GeneChip® data were in accordance with the published expression patterns, thus validating the combined approaches of laser capture microdissection and microarray technology to produce data of reasonable quality, albeit with the caveat that some technical variability is apparent at low-abundance levels of mRNA abundance.

4.8.3 promoter::GUS analysis of differentially regulated putative transcription factors

Comparison of my GeneChip® data to published expression patterns for known embryonic genes provided a basis of validation to the data. However, the number of genes available with well-characterized expression patterns is limited. Therefore, in order to provide an additional level of validation and to

initiate further research based on the data produced, a number of promoter::GUS constructs were generated for genes predicted to encode putative transcription factors that were hypothesized to exhibit differential spatial expression, based on the Affymetrix data obtained. Five promoter::GUS fusions were successfully constructed and introduced into transgenic plants. At the torpedo-stage of embryogenesis the 5 genes were predicted to be expressed either constitutively (At2g31510/R and At5g45600/T), predominantly in the cotyledons (At5g14610/C and At5g50810/E) or predominantly in the root (At1g78160/B).

The promoter::GUS activities for the predicted constitutively expressed constructs R and T corresponded with the GeneChip® data. Construct R displayed a higher intensity of staining in the root compared to the cotyledons which was predicted from the very high root signal value seen in the heart-stage GeneChip® data obtained from Casson *et al.* (2005). Construct T also had a very high signal value in the heart-stage root, although the expression at the torpedo-stage was entirely constitutive as predicted. Both constructs constitutive expression patterns were maintained in the seedling.

The promoter::GUS activities for the predicted cotyledon expressed constructs C and E also correlated well with the GeneChip® data. Construct C exhibited a clear apical localisation in the torpedo-stage embryo and then exhibited a very specific expression pattern to the stomatal guard cells in the seedling. Construct E displayed a highly variable expression profile through embryogenesis and into the seedling, which provides an important insight into the success of the technique in defining temporal and spatial expression. In the heart-stage data of Casson *et al.* (2005), At5g50810 expression is predicted to be predominantly in the root, contrasting markedly with the predominantly cotyledon expression predicted in my torpedo-stage data. However, while expression in a late-heart/early-torpedo-stage embryo appears largely constitutive, by the mid-torpedo-stage there is no expression detectable in the root correlating with the predicted expression pattern. To complete the temporal and spatial expression changes time-course, at the 7dpg seedling stage of development the promoter::GUS activity pattern is once again constitutive but with weak staining in the cotyledon vasculature compared to very intense staining in the root. The

marked changes in expression pattern pre- and post- the torpedo-stage for this gene show a distinct spatial and temporal pattern, which my data accurately represent.

Construct B was predicted to show promoter::GUS activity predominantly in the root. Preliminary analysis was conducted and a single line displaying promoter::GUS activity was isolated. A single heart-stage embryo was located that showed promoter::GUS activity, with the pattern appearing to be constitutive. Further characterisation is required to confirm the expression pattern observed in the torpedo-stage embryo. No staining was detectable in the seedlings of a number of lines analysed (not including the line displaying promoter::GUS activity in the embryo), and therefore also requires further investigation.

4.8.4 Functional analysis of differentially regulated putative transcription factors

Genome-wide expression studies are a very powerful tool in dissecting out genes of putative importance in a spatial and temporal dimension. However, given that the relative mRNA abundance is not an absolute measure of the activity of a gene product *in situ*, they cannot be taken to represent the complete picture. Therefore a reverse genetics approach must be adopted, taking genes of interest and analysing the phenotype of transgenic plants in which their expression is disrupted.

Insertional mutagenesis by T-DNA integration leads to the knockout of gene expression, providing a direct route to uncovering the gene products *in situ* function (see Section 1.1.1.4). In addition to its chemical and physical stability post-integration the T-DNA also acts as a marker for ease of mutant identification.

The SALK T-DNA insertion lines analysed in this work were predicted to represent null mutations, producing phenotypes from which biological function could be predicted. However, of the lines analysed no phenotype could be

attributed to the knockout mutants. This situation does not appear to be uncommon, with many SALK knockout mutants not displaying an aberrant phenotype (Casson, S.A. and Chilley, P. personal communication). Indeed, amongst a collection of SALK lines expected to result in null phenotypes for *EMB* genes of unknown function a range of possible outcomes were observed including: failure to produce plants with the expected insertion; contradictory results; and plants with a confirmed insert but no associated phenotype (Tzafrir *et al.*, 2004). Possible explanations for this are that the phenotype is conditional, that functional redundancy exists between members of a gene family or that the T-DNA insertion has not significantly disrupted gene function.

Krysan *et al.* (1996) identified T-DNA insertion lines for 17 genes involved in signal transduction and ion transport. Under normal growth conditions none of these lines displayed an aberrant phenotype. However, when grown under specific a condition, which, in the case of the AKT1 potassium channel was growth on media with a potassium concentration of 100 μM or less, a mutant phenotype was observed (Hirsch *et al.*, 1998; Krysan *et al.*, 1999). Such conditional phenotypes may be the result of evolution amongst gene family members in order to function under specific physiological conditions, and thus the phenotype is only observed when the knockout plant is placed under physiological conditions in which the gene product is required (Hirsch *et al.*, 1998). Conditional phenotypes for a gene of unknown function can be tested for using a range of physiologically challenging conditions under which wild-type plants would be expected to survive.

An alternative explanation for the lack of an identifiable phenotype in knockout plants is that of functional redundancy existing amongst the members of a gene family. Of the putative transcription factors analysed in this work several belong to recognised families with multiple members: At1g63900 and At2g31510 belong to the C3H family which comprises 165 members (http://arabidopsis.med.ohio-state.edu/AtTFDB/); At2g45050 belongs to the C2C2-GATA family which comprises 30 members (http://arabidopsis.med.ohio-state.edu/AtTFDB/); and At5g14610 belongs to the DEAD box RNA helicase family which comprises 53 members (Aubourg *et al.*, 1999; Boudet *et al.*, 2001). The remainder have not been assigned to recognised families, although

At2g25420 harbours a WD-40 domain which would give it homology to ~400 other genes (http://www.ebi.ac.uk/interpro/IEntry?ac=IPR001680). Examples of partial or full functional redundancy have been uncovered amongst gene families of importance at many different stage of development.

The three NAC domain transcription factors, *CUP-SHAPED COTYLEDON1* (*CUC1*), *CUC2* and *CUC3*, act in a partially functionally redundant fashion to suppress fusion of the cotyledon primordia, and regulate the formation of the embryonic SAM by the induction of *STM* expression (Aida *et al.*, 1997, 1999; Takada *et al.*, 2001; Vroemen *et al.*, 2003).

Similarly, the four members of the PIN family of auxin efflux carriers, which are expressed in the *Arabidopsis* embryo, appear to act redundantly with only minor phenotypic aberrations observed in single mutants, this is despite the distinct temporal and spatial expression patterns observed between family members precluding a direct substitution of one PIN protein for another. When combined in double, triple or quadruple mutant combinations, significantly more severe defects are observed (Jenik and Barton, 2005). Recent research suggests that in mutant plants the remaining functional PIN proteins undergo a redistribution to counteract the loss of the other PIN proteins (Blilou *et al.*, 2005).

An example of complete functional redundancy is that of the MADS domain protein-encoding *CAULIFLOWER (CAL)* gene of *Arabidopsis,* which as a single mutant appears phenotypically normal. Contrastingly, the highly similar MADS domain encoding *APETALA1 (AP1)* gene, when mutated results in defective floral meristem and floral organ specification. However the *cauliflower* phenotype is uncovered in *apetala1 cauliflower* double mutants which exhibit a greatly enhanced abnormal phenotype including developmental arrest at the inflorescence meristem stage, thus demonstrating *CAULIFLOWER* is functionally redundant to *APETALA1* (Mandel *et al.*, 1992; Bowman *et al.*, 1993; Kempin *et al.*, 1995).

Hua and Meyerowitz (1998) described functional redundancy in a family of putative ethylene receptors, which showed between 57% and 79% sequence similarity to one another. Single mutants displayed no ethylene response

defects, while a quadruple mutant manifested a constitutive ethylene triple response. Triple and double mutants displayed progressively diminished ethylene response defects demonstrating overlapping roles in ethylene perception.

To test for such functional redundancy, knockout mutants must be obtained in multiple members of a gene family and double, triple and quadruple mutants created from these to assess for any additional phenotypic changes as were seen in the examples above.

4.8.5 Expression patterns of known *EMB* genes

The *EMB* gene collection represents the estimated 500-1000 genes in *Arabidopsis* which when mutated produce an embryo-defective phenotype (Franzmann *et al.*, 1995; McElver *et al.*, 2001). These genes perform essential functions in embryogenesis and are required for viability under normal conditions. Of the collection of 220 *EMB* genes described by Tzafrir *et al.* (2004), ~96% are deemed to be expressed in the torpedo-stage embryo when a signal value cut-off of 40 is imposed, the remaining 4% are either expressed at a different stage of embryogenesis (Casson *et al.*, 2005) or are predicted to expressed in a tissue region not analysed in this study (e.g. the hypocotyl). This % of *EMB* genes predicted to be expressed at each signal cut-off value is higher for the torpedo-stage than that observed for the globular- and heart-stage data (Casson *et al.*, 2005), this could be due to a number of *EMB* genes being important at later stages of embryogenesis as illustrated by the greater number of terminal phenotypes observed at later (cotyledonary) stages of embryogenesis than early (pre-globular) stages (Tzarfrir *et al.*, 2004).

As many of the *EMB* genes appear to be of low-abundance, as indicated by low signal values in the GeneChip® data, care must be taken when assessing some of the fold-changes observed between different embryonic regions as the difference in background hybridisation level can create somewhat artificial results. For instance, a background hybridisation signal value of 5 would indicate a significantly larger fold change compared to 100 than would a

background hybridisation signal value of 30, although both of these values could be considered to represent a non-expressed gene. However, a number of *EMB* genes appeared to show distinct spatial patterns of differential expression.

The serine-threonine protein kinase *PINOID*, described previously, is highly up-regulated (15-fold more abundant) in the cotyledon compared to the root correlating with the defective cotyledons morphology observed in the *pinoid* mutant (Christensen *et al.*, 2000).

In the root and meristem two members of the *TITAN* (*TTN)* family show up-regulation compared to the cotyledons (Liu and Meinke, 1998). *TTN6* encodes a deubiquitinating enzyme, and was shown to be 4-fold more abundant in the root compared to the cotyledon where there is likely no expression. Embryonic cells in the *ttn6* mutant were rounded and disorganised, culminating in developmental arrest at the globular-stage, in addition to a disruption of endosperm cellularisation (Doelling *et al.*, 2001; Tzafrir *et al.*, 2002). *TTN6* did not emerge as significantly spatially distributed in the analysis of *EMB* gene differential expression conducted by Casson *et al.* (2005) on globular and heart-stage embryos. In light of the developmental arrest observed at the globular-stage of embryogenesis in *ttn6* mutants, perhaps this suggests that constitutive expression is required at these critical stages early of embryogenesis with a contrasting spatial expression pattern required for functions during late embryogenesis. *TTN9* appears to be expressed in both the root and shoot apical meristem region, but not in the cotyledons. The *ttn9* showed significant enlargement of endosperm nucleoli and developmental arrest of embryogenesis at a very early stage, reaching a maximum of four cells and resembling the *ttn7* and *ttn8* mutant phenotypes, which result from the knock-out of cohesin (Liu *et al.*, 2002; Tzafrir *et al.*, 2002).

SHOOT MERISTEMLESS (*STM*), a *KNOTTED1*-like homeobox *(KNOX)* gene is found to be ~62-fold and ~19-fold more abundant in the shoot apical meristem than in the cotyledons and root respectively, as cautioned earlier these fold-changes are slightly mis-leading and represent slightly different background levels of hybridisation, *STM* doesn't appear to be expressed in either the cotyledons or root, as confirmed by *in situ* hybridisation (Long *et al.*,

1996). This expression pattern is complemented by the *stm* mutant phenotype in which shoot meristem development is blocked from the torpedo-stage of embryogenesis onwards (Barton and Poethig, 1993).

The validation of my GeneChip® data can be considered a success, with the majority of expression patterns analysed found to be highly comparable. With some degree of confidence attached, the data can now be subjected to high-level analysis in an attempt to provide a more global perspective on the transcriptional profiling of plant embryogenesis.

Chapter 5

Results:

Transcriptional Profiling of Plant Embryogenesis

5.0 Transcriptional profiling of plant embryogenesis

5.1 Introduction and objectives

The previous chapter demonstrated the application of microarray technology to LCM embryonic material. A preliminary analysis of the Affymetrix MAS 5.0 normalised data obtained was performed using Excel to rank genes according to fold-changes observed between cotyledonary and root-derived samples. The microarray data were validated using a range of approaches including literature searches and promoter::GUS analysis.

The work described in this chapter aims to address the questions of gene regulation driving apical-basal pattern formation on a more global level through the use of high-level computational analysis. GeneSpring version 7.2 (Silicon Genetics, Redwood City, CA), a commercially available visualisation and analysis package widely regarded as the gold standard for expression data analysis was utilised to achieve these aims. To accomplish this analysis a number of GeneChip® data sets were utilised. The GeneChip® data obtained from LCM of torpedo stage embryos were analysed alongside GeneChip® data for globular and heart stage embryos produced by Casson *et al.* (2005), and the objective was to show the changing gene expression patterns through the embryonic time-course in both the apical and basal domains.

In addition to analysing gene expression changes over the embryonic time-course, a comparison was made between the apical and basal region at each developmental stage in an attempt to uncover genes of significance in the maintenance of apical-basal polarity. Significant apical and basal genes from each stage were compared in an attempt to identify common genes of importance. For an additional comparative time-point GeneChip® data obtained from the non-embryonic cotyledon and root tissue of a seedling, 7 days after germination, were employed (AtGenExpress Consortium, http://www.affymetrix.arabidopsis.info).

The comparisons undertaken are shown in Figure 5.1

Figure 5.1 Comparative regions for GeneSpring analysis.

1. Globular-stage embryo
2. Heart-stage embryo
3. Torpedo-stage embryo

Apical/Cotyledonary region

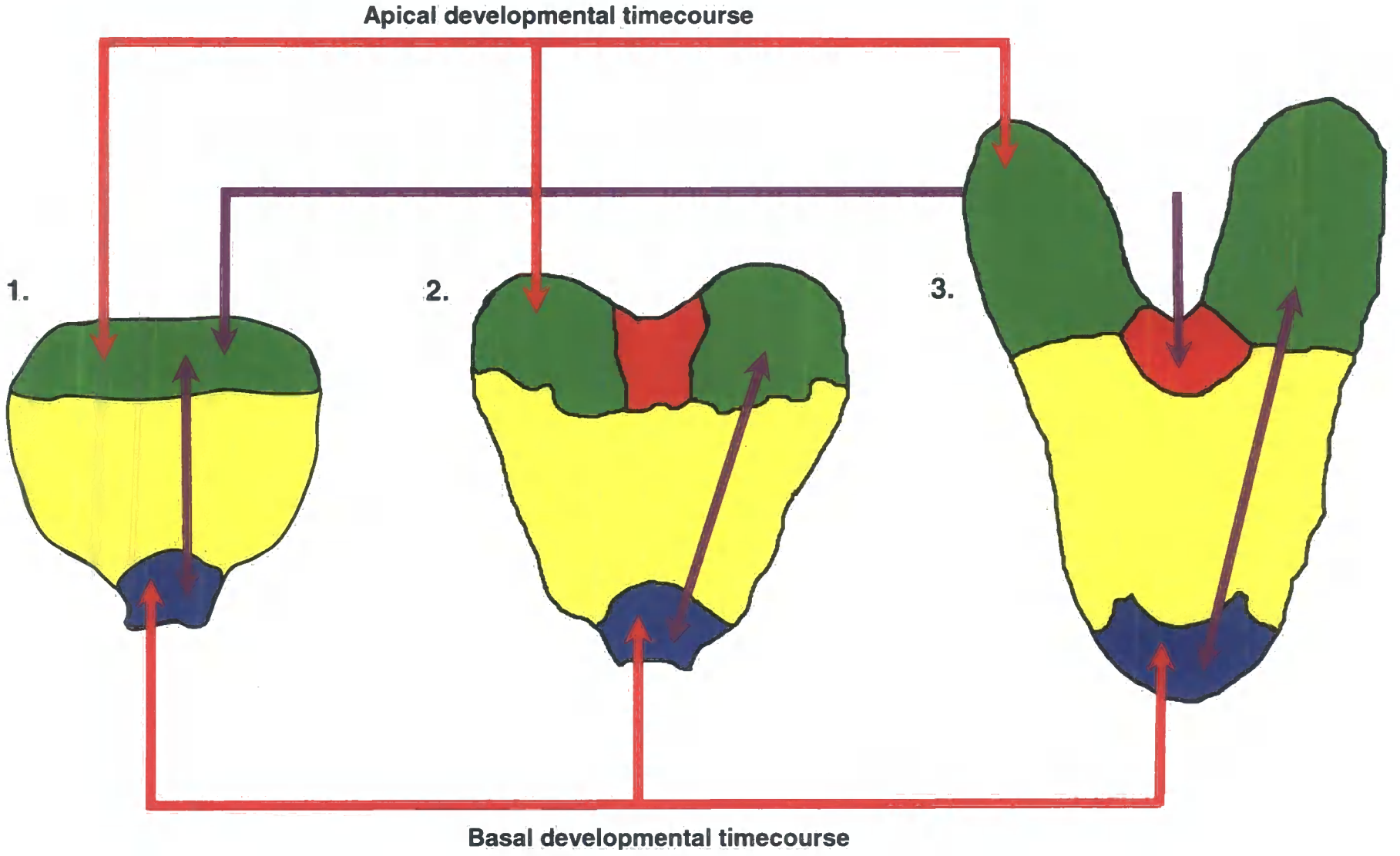
Presumptive shoot apical meristem

Hypocotyl region

Basal/Root region

Pairwise comparison of embryonic regions

Developmental timecourse comparison


Apical developmental timecourse
1.
2.
3.
Basal developmental timecourse

5.2 Data handling and normalisation

The raw GeneChip® data was normalised using robust multi array average (RMA), a log scale measurement of expression developed by Irizarry *et al.* (2003), based on PM values only rather than PM and MM.

For individual experiments a second step of per gene normalisation was imposed. For experiments involving the analysis of more than two samples (i.e. a time-course) each expression value was normalised by dividing it by the median of all expression values, as shown in the formula below.

$$\frac{\textbf{(signal strength of gene A in sample X)}}{\textbf{(median of every measurement taken for gene A throughout your experiment)}}$$

For pair-wise comparisons all expression values were normalised to a control sample (i.e. the basal sample in an apical-basal comparison), such that each gene was divided by the intensity of that gene in the control sample, as shown below.

$$\frac{\textbf{(signal strength of gene A in sample X)}}{\textbf{(signal strength of gene A in the control sample)}}$$

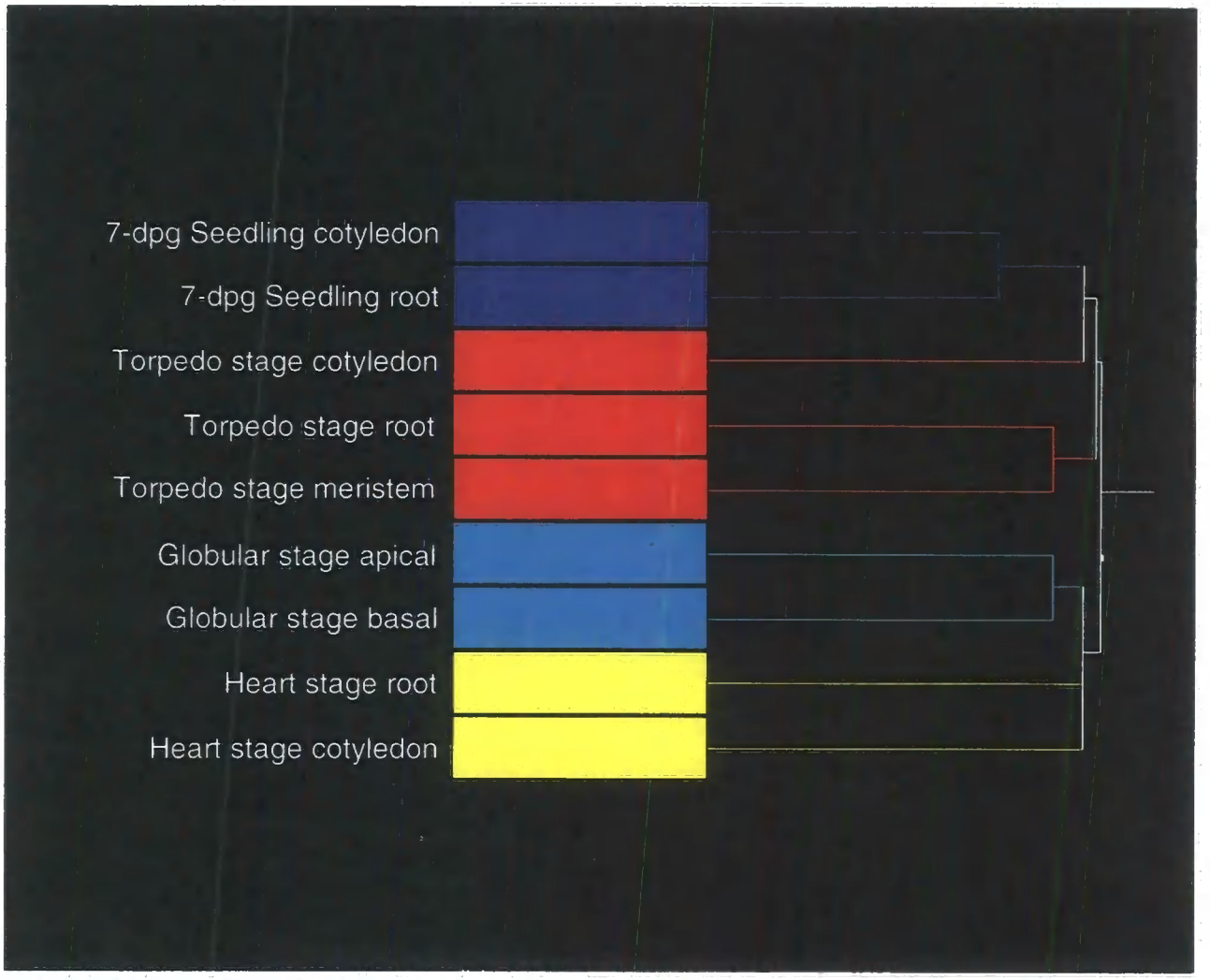
(Formulae taken from the GeneSpring version 7.0 manual).

5.3 Relatedness of embryonic developmental stages

Each of the embryonic stages under investigation is defined by a distinct phenotype, and it can therefore be hypothesised that each developmental stage has its own distinct transcriptional profile underlying its phenotype. It was decided to conduct an experiment to test this hypothesis and to assess the

Figure 5.2 Condition tree cluster analysis of developmental stages.

All the samples normalised together to a per gene median value. Clustering analysis performed using condition tree clustering on all samples. Similarity measured using Spearman correlation (GeneSpring version 7.2).


7-dpg Seedling cotyledon
7-dpg Seedling root
Torpedo stage cotyledon
Torpedo stage root
Torpedo stage meristem
Globular stage apical
Globular stage basal
Heart stage root
Heart stage cotyledon

degree of similarity shown between the transcriptional profiles of the different samples under investigation.

All the samples were initially normalised together to a per gene median value. Clustering analysis was then performed using condition tree clustering on all samples. Similarity was measured using Spearman correlation (GeneSpring version 7.2).

Cluster analysis of the entire transcriptional profile of roots (basal) and cotyledons (apical) of the developmental stages globular, heart, torpedo and seedling, and torpedo stage SAM, revealed that the transcriptional profiles of the apical regions are more closely related to their respective developmental stage basal region than they are to the other apical regions (Figure 5.2). This supports the hypothesis that each developmental stage has its own distinct transcriptional profile and suggests that genes with specifically apically and basally localised expression patterns contribute only a minority of the overall profile. The condition clustering analysis also makes a clear separation between the early embryonic globular and heart stages, and the later torpedo stage, which is calculated to be closer to the seedling in terms of its transcriptional profile. Interestingly, the torpedo-stage SAM clusters closer to the torpedo-stage root than it does to the torpedo-stage cotyledon indicating a clear transcriptional difference between these adjacent regions.

5.4 Transcriptional changes along an embryonic developmental time-course

Having established that the transcriptional profile of an apical region at a particular developmental stage has a greater similarity to its respective basal region than to apical regions of other developmental stages, it was decided to look in greater detail at the changes taking place between stages. A separate analysis was undertaken for both the apical (cotyledon) regions and the basal (root) regions. While not directly targeting genes of potential importance in apical-basal polarity, it was hoped that such an analysis would provide an insight into potential differences in the functional gene classes of importance in

the different regions. A clustal analysis was also undertaken with the aim of elucidating potentially important groups of genes with similar expression profiles across the three stages.

5.4.1 Apical developmental time-course

All the apical (cotyledon) samples were normalised together to a per gene median value. A graphical view of the normalised data is shown in Figure 5.3, with the expression value of each gene plotted on a log scale against the developmental time-course. As can clearly be seen the three developmental stages all have very distinct outlines around a core of genes centred on the default expression value of 1. The globular stage displays the highest degree of variability with a number of genes of massively high and low expression values. At the heart stage there appears to be a similar general spread of high and low expression values but without the extreme outliers present at the globular stage. In comparison the torpedo stage data shows a much smaller range of expression levels. Figure 5.3 includes data for all the genes present on the GeneChip®, not all of which are likely to be expressed at a given stage, additionally no statistical significance has been attributed to the expression values displayed, however despite these limitations it provides a useful measure of the potential differences which could be present between the stages.

To assess changes in gene expression patterns on a functional level the data was filtered by significance using a Student's t-test with a maximum confidence level of 5% for genes whose expression was significantly different from a value of one. 1872 genes satisfied this criterion in at least one of the three developmental stages. Further filtering was accomplished by calculating a fold-change between the expression values at different developmental stages. Comparisons were made between globular and heart; globular and torpedo and heart and torpedo. In each case the 100 most up-regulated genes passing the significance filter were selected. These genes were assigned functional annotation using information from http://mips.gsf.de/proj/thal/db/ and Figure 5.4 displays the functional classifications of genes up-regulated between

developmental stages. Comparing the heart and torpedo stages to the globular stage, it is interesting to note the up-regulation of genes involved in energy production, for example the photosystem, which comprise 21% of the up-regulated heart stage genes and 31% of the up-regulated torpedo stage genes. Other significant functional groups represented are metabolism (19% at heart stage, 14% at torpedo stage), cellular communication/signal transduction (7% at heart stage) and transcription (7% at heart stage, 8% at torpedo stage). Comparing the torpedo stage to the heart stage, genes related to the production of energy form 20% of the up-regulated genes, and once again these are heavily biased towards the photosystem. Also of note is that 15% of the up-regulated genes are involved in protein synthesis, perhaps reflecting a change in emphasis as the embryo progresses towards late embryogenesis. Metabolic genes also comprise 14% of the total. Through all the functional comparisons those genes of unknown function represented between 22% and 31% of the total.

K-means clustal analysis was performed on the 1872 genes satisfying the significance criteria, using Pearson correlation (GeneSpring version 7.2). The user defines the maximum number of clusters formed; in this case 5, 10 and 15 cluster experiments were performed. The 10-cluster analysis is shown in Figure 5.5. It was found that approximately 7 distinct expression patterns are present within the sampled genes; these trends are illustrated diagrammatically in Figure 5.6.

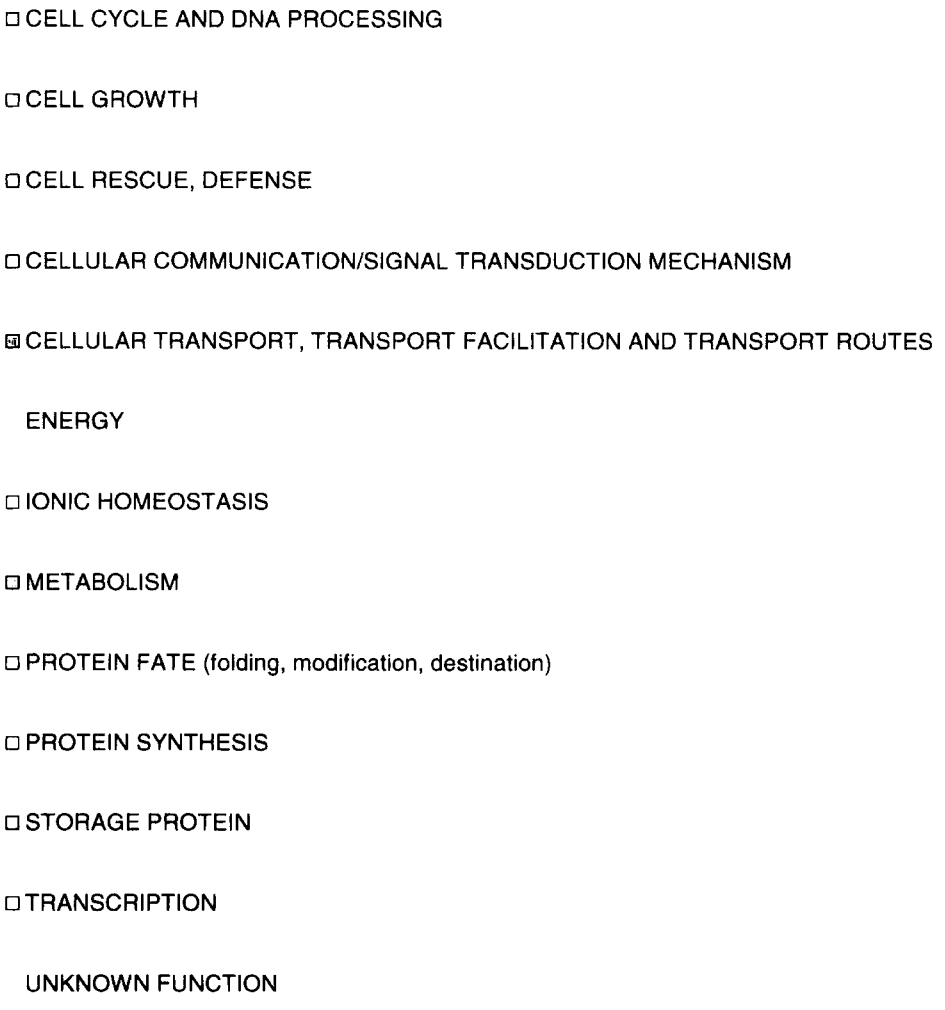
An analysis was carried out to determine whether any of the clusters obtained were specifically enriched for particular families of predicted transcription factors or receptor kinases. A database of approximately 1400 predicted transcription factors and receptor kinases (Davuluri *et al.*, 2003; Shiu and Bleecker, 2003; http://arabidopsis.med.ohio-state.edu/AtTFDB/) was used to probe the clusters. Due to the low numbers of transcription factor family members present in the filtered gene list no valid statistical significance could be attributed to the numbers appearing in individual clusters. Despite this limitation no individual cluster showed any notable enrichment for particular families of transcription factors.

Figure 5.3 Apical developmental timecourse.

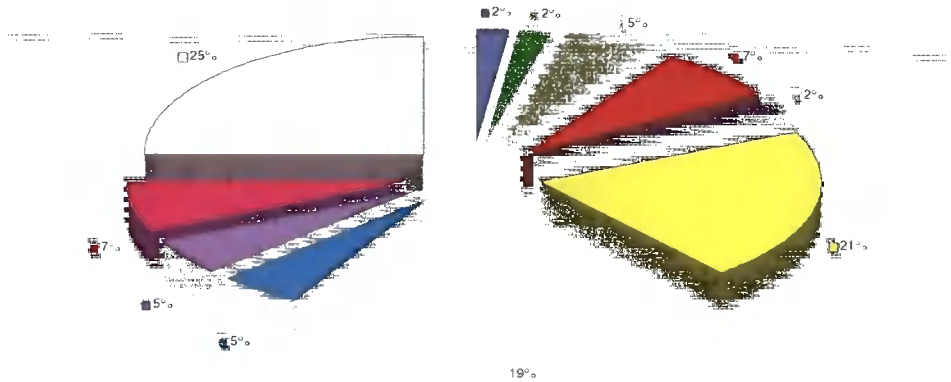
All the apical (cotyledon) samples were normalised together to a per gene median value. The expression value of each gene is plotted on a log scale against the developmental stage.

Up-regulated Genes
Unchanged
Down-regulated Genes
Normalised Intensity (log scale)
100
10
1
0.1
Globular
Heart
Torpedo
Developmental Stage

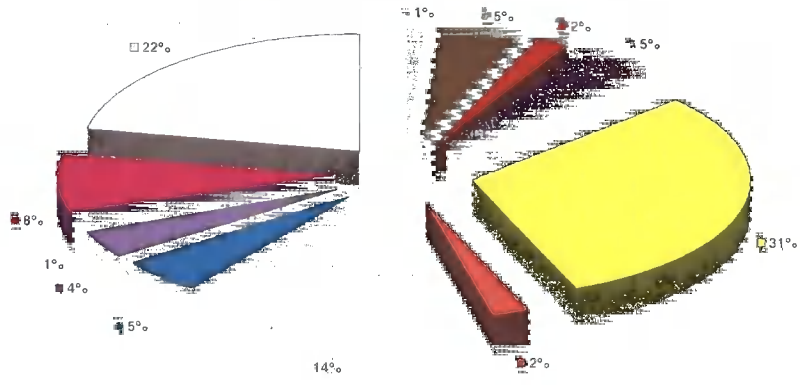
Figure 5.4 Functional annotations of the 100 most up-regulated genes (passing the significance filter: $p<0.05$ in at least one stage) between developmental stages on the apical developmental timecourse.



Functional classification of genes up-regulated at the heart stage compared to the globular stage



Functional classification of genes up-regulated at the torpedo stage compared to the globular stage



Functional classification of genes up-regulated at the torpedo stage compared to the heart stage

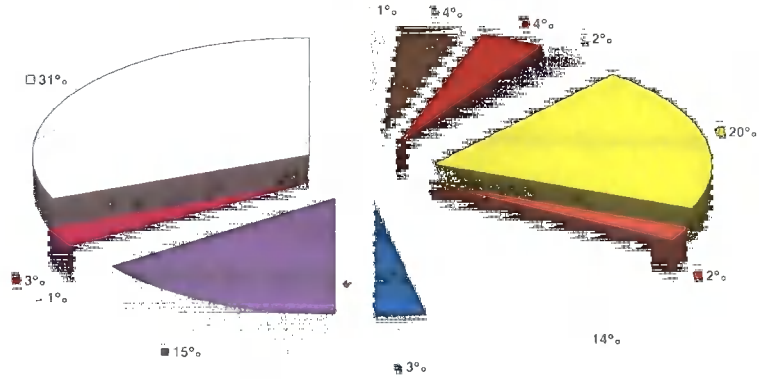
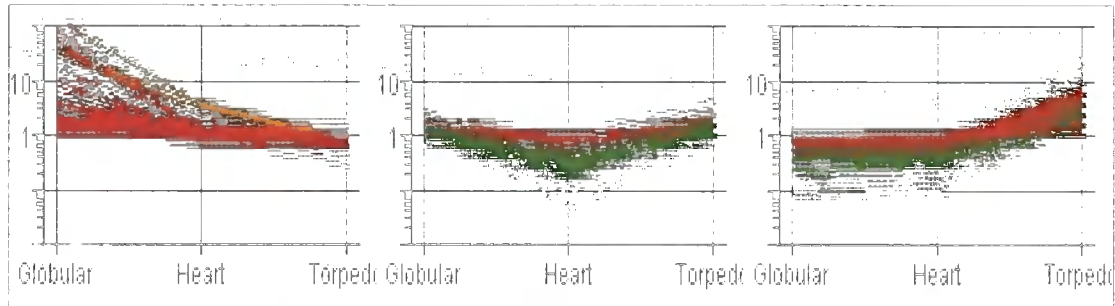


Figure 5.5 10-cluster K-means analysis performed on the apical developmental timecourse.

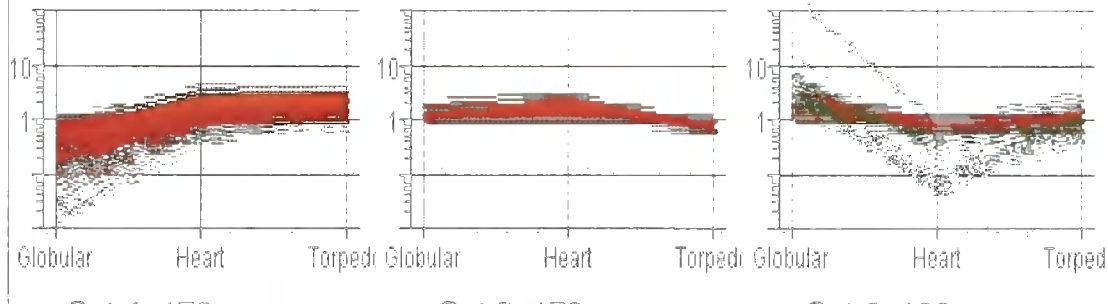
All the samples normalised together to a per gene median value. 1872 genes satisfied the significance criteria ($p<0.05$ in at least one of the three stages). Clustering analysis performed using K-means clustering on all samples. Similarity measured using Pearson correlation (GeneSpring version 7.2).



Set 1: 276 genes

Set 2: 103 genes

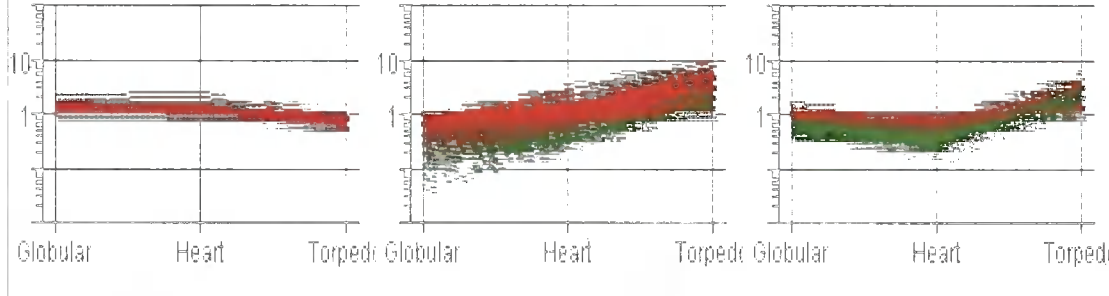
Set 3: 261 genes



Set 4: 178 genes

Set 5: 178 genes

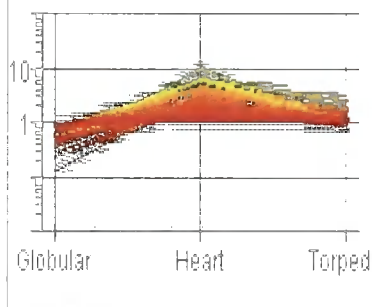
Set 6: 132 genes



Set 7: 282 genes

Set 8: 202 genes

Set 9: 134 genes



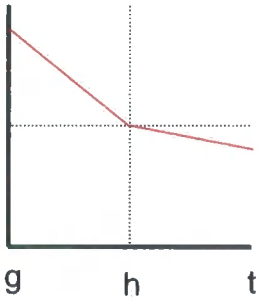
Set 10: 126 genes

Figure 5.6 Distinct temporal expression patterns emerging from K-means clustal analysis of an apical developmental series.

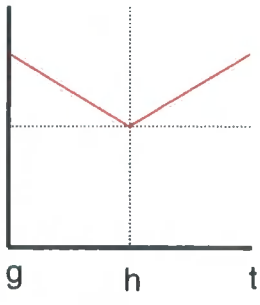
Horizontal axis: Developmental stage (globular, heart and torpedo)
Vertical axis: Expression level

K-means clustal analysis was performed on the 1872 genes satisfying the significance criteria, using Pearson correlation (GeneSpring version 7.2). It was found that approximately 7 distinct expression patterns trends are present within the sampled genes.

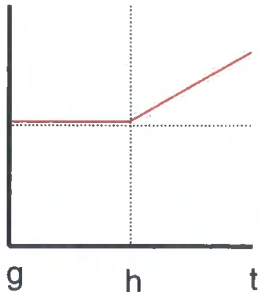
1.



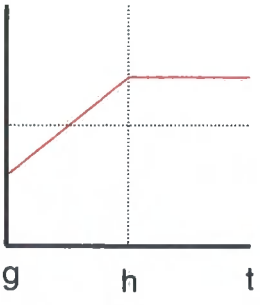
2.



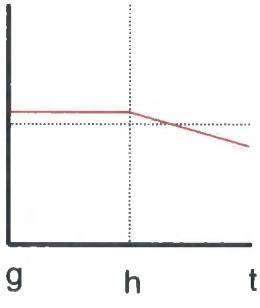
3.



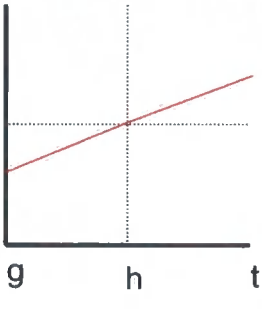
4.



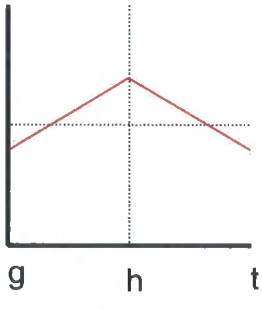
5.



6.



7.



5.4.2 Basal developmental time-course

As with the apical time course, all the basal (root) samples were normalised together to a per gene median value. A graphical view of the normalised data is shown in Figure 5.7, with the expression value of each gene plotted on a log scale against the developmental time-course. Once again each stage has a distinct profile, which in many ways are similar to those observed for the apical region. However in contrast there appears to be a considerably more compact profile centred around the expression value range 1-3, with a reduced number that have relatively extreme expression values. As before, this Figure represents the entire GeneChip® data set with all the attached limitations, however, a clear similarity of expression profile is shown between the developmental stage datasets and between the overall apical and basal time-course profiles.

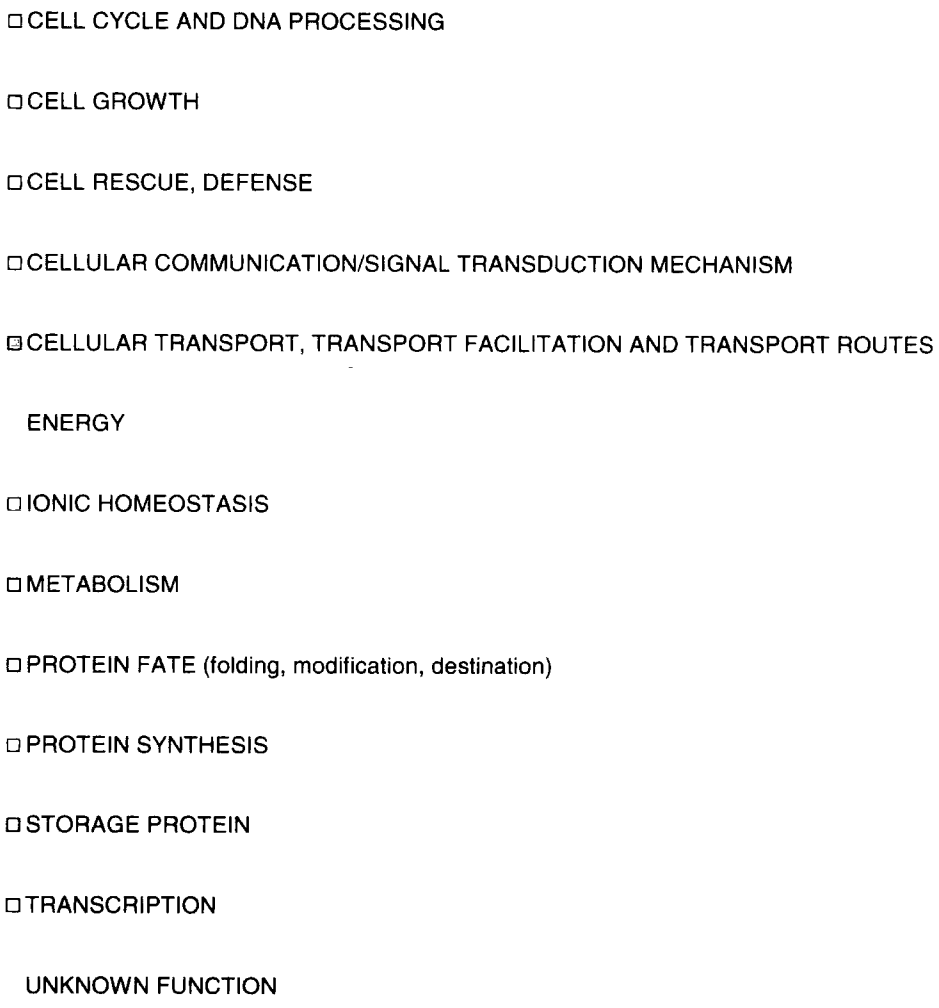
As for the apical time-course analysis, the data were filtered by significance using a Student's t-test with a maximum confidence level of 5% for genes whose expression was significantly different from one. For the basal developmental time-course 1226 genes satisfied this criterion in at least one of the three developmental stages. Further filtering was accomplished by calculating a fold-change between the expression values at different developmental stages. Comparisons were made between globular and heart; globular and torpedo; and heart and torpedo. In each case the most up-regulated 100 genes passing the significance filter were selected. These genes were assigned functional annotation using information from http://mips.gsf.de/proj/thal/db/ and Figure 5.8 displays the functional classifications of genes up-regulated between developmental stages. Comparing the heart stage to the globular stage, four main functional groups are represented: metabolism (27%), energy (9%), protein synthesis (8%) and transcription (8%). A different profile is observed in a comparison of the torpedo stage and globular stage with many functional classes being represented at around 5% of the total with metabolism (16%), energy (8%), transcription (9%) and cell growth (7%) being slightly above this level. Comparing the torpedo stage to the heart stage, five main functional groups are represented: metabolism (15%), cell growth (14%), cell rescue/disease (8%), transcription

Figure 5.7 Basal developmental timecourse.

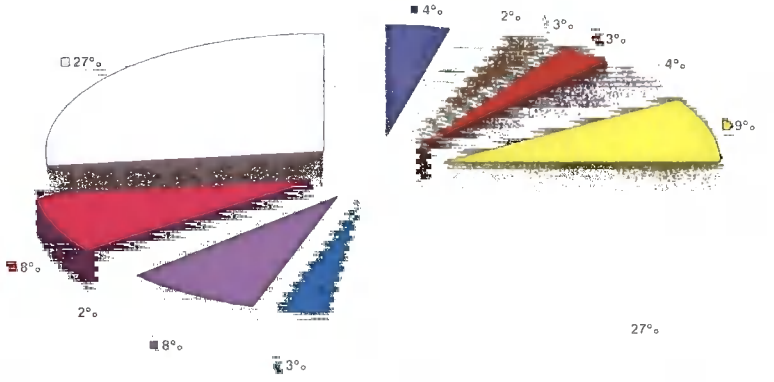
All the basal (root) samples were normalised together to a per gene median value. The expression value of each gene is plotted on a log scale against the developmental stage.

Up-regulated Genes
Unchanged
Down-regulated Genes
Normalised Intensity (log scale)
100
10
1
0.1
Globular
Heart
Torpedo
Developmental Stage

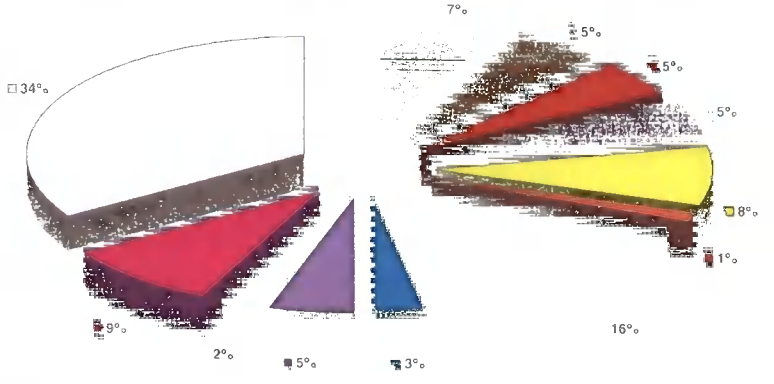
Figure 5.8 Functional annotations of the 100 most up-regulated genes (passing the significance filter: $p<0.05$ in at least one stage) between developmental stages on the basal developmental timecourse.



Functional classification of genes up-regulated at the heart stage compared to the globular stage



Functional classification of genes up-regulated at the torpedo stage compared to the globular stage



Functional classification of genes up-regulated at the torpedo stage compared to the heart stage

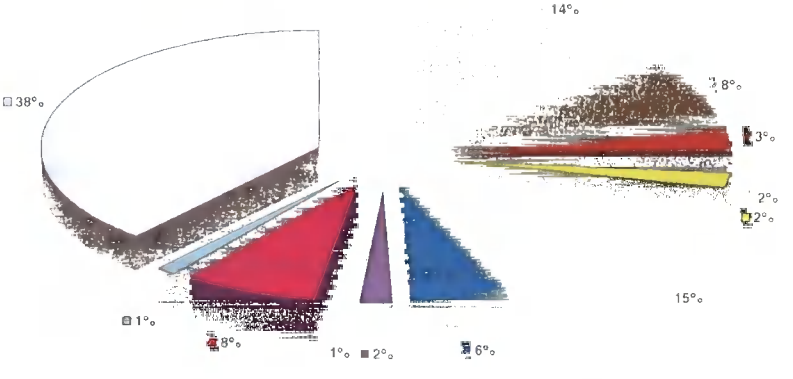


Figure 5.9 10-cluster K-means analysis performed on the basal developmental timecourse

All the samples normalised together to a per gene median value. 1226 genes satisfied the significance criteria ($p<0.05$ in at least one of the three stages). Clustering analysis performed using K-means clustering on all samples. Similarity measured using Pearson correlation (GeneSpring version 7.2).

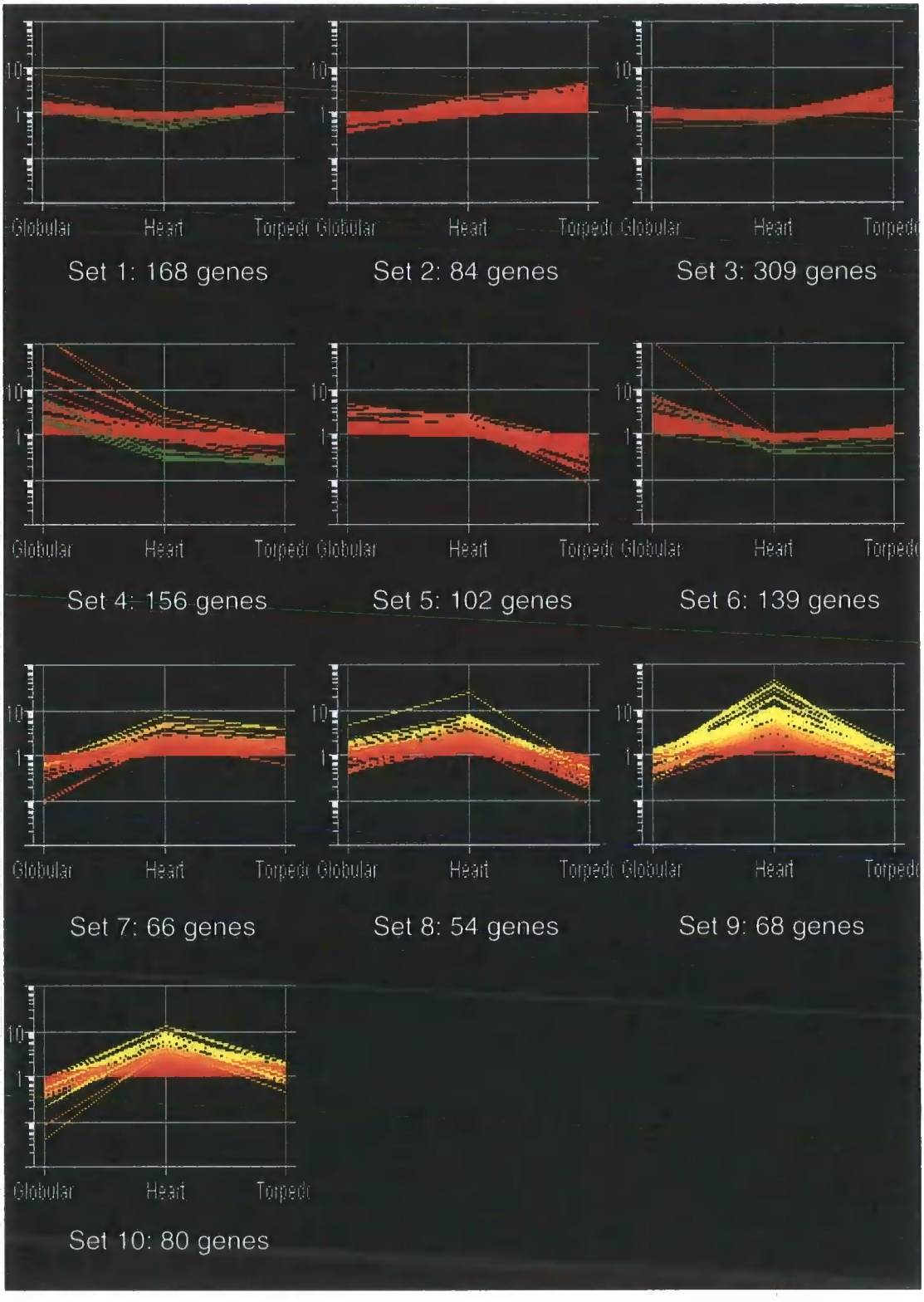
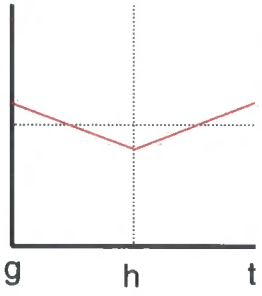

10
1
Globular
Heart
Torpedo
Set 1: 168 genes
Set 2: 84 genes
Set 3: 309 genes
Set 4: 156 genes
Set 5: 102 genes
Set 6: 139 genes
Set 7: 66 genes
Set 8: 54 genes
Set 9: 68 genes
Set 10: 80 genes

Figure 5.10 Distinct temporal expression patterns emerging from K-means clustal analysis of a basal developmental series.

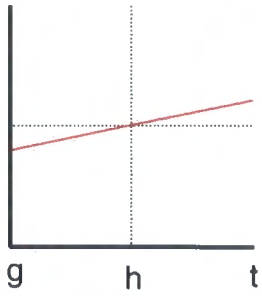
Horizontal axis: Developmental stage (globular, heart and torpedo)
Vertical axis: Expression level

K-means clustal analysis was performed on the 1226 genes satisfying the significance criteria, using Pearson correlation (GeneSpring version 7.2). It was found that approximately 7 distinct expression patterns trands are present within the sampled genes.

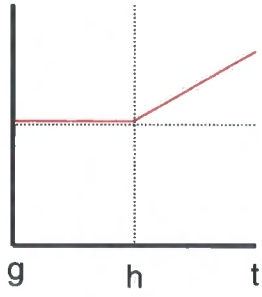
1.



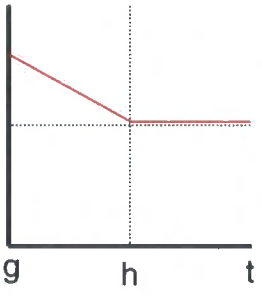
2.



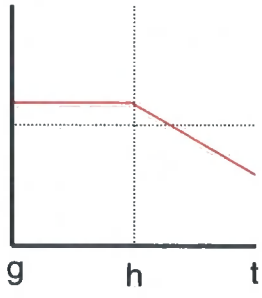
3.



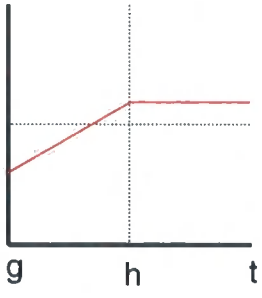
4.



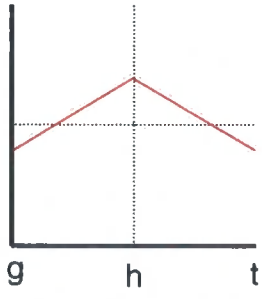
5.



6.



7.



(8%) and protein fate (6%). As with the apical timecourse functional analysis a change in emphasis appears to occur between the heart and the torpedo stage possibly reflecting the approach of late embryogenesis. The fraction accounted for by genes with an unknown function is even higher than that of the apical region comprising between 27 and 38% of the total.

K-means clustal analysis was performed on the 1226 genes satisfying the significance criteria, using Pearson correlation (GeneSpring version 7.2). The user defines the maximum number of clusters formed; in this case 5, 10 and 15 cluster experiments were performed. The 10-cluster analysis is shown in Figure 5.9. It was found that approximately 7 distinct expression patterns are present within the sampled genes; these trends are illustrated diagrammatically in Figure 5.10.

An analysis was carried out to determine whether any of the clusters obtained were specifically enriched for particular families of predicted transcription factors or receptor kinases. A database of approximately 1400 predicted transcription factors and receptor kinases (Davuluri *et al.*, 2003; Shiu and Bleecker, 2003; http://arabidopsis.med.ohio-state.edu/AtTFDB/) was used to probe the clusters. Due to the low numbers of transcription factor family members present in the filtered gene list no valid statistical significance could be attributed to the numbers appearing in individual clusters. Despite this limitation no individual cluster showed any notable enrichment for particular families of transcription factors.

5.6 Tissue comparisons

The work described in the previous sections has shown that individual developmental stages have their own distinct transcriptional profiles, highlighted by the apical regions clustering more closely to the basal regions of the same developmental stage than to the apical regions of other stages (see Section 5.3). However, an analysis of genes up-regulated along an apical and basal timecourse showed clear differences in the relative prevalence of certain functional classes and thus indicated distinct transcriptional profiles for the two

tissue regions (see Sections 5.4 and 5.5). To further this study into the transcriptional profiles of apical and basal regions, it was decided to specifically compare the apical and basal region at each stage along the developmental time-course. In addition a comparison was also made between GeneChip® data for the cotyledon and root tissue of a 7dpg seedling, produced by the AtGenExpress Consortium, and provided by NASCArrays at http://www.affymetrix.arabidopsis.info. As well as analysing the most differentially expressed genes between the regions of a particular developmental stage, a study was also carried out to compare the up-regulated genes of each region/developmental stage to assess the degree of overlap and therefore the degree to which these genes were specifically apical or basal throughout development.

5.6.1 Globular apical versus basal comparison

The globular apical and basal samples were normalised together to a control sample, which in this case was the basal sample, therefore genes in the apical sample had an expression value of either >1 (up-regulated), <1 (down-regulated) or 1 (identical expression in both tissue samples). The resulting data was then filtered by significance using a Student's t-test with a maximum confidence level of 5% for genes whose expression was significantly different from a value of one. To correct for the occurrence of false positives a Benjamini and Hochberg false discovery rate, multiple testing corrections was used to adjust the p-values (Benjamini and Hochberg, 1995; GeneSpring version 7.2).

585 genes satisfied these criteria, and were sorted into those up-regulated and down-regulated in the apical region compared to the basal region. Due to the comparative nature of the normalisation these gene lists also corresponded to down-regulated and up-regulated genes in the basal region compared to the apical region. Therefore, these sorted gene lists correspond to apical and basal up-regulated genes, albeit at this stage including genes with potentially very little difference in expression value between the two tissue types.

Table 5.1 **The ten most differentially expressed significant genes up-regulated in the apical region of the globular stage embryo compared to the basal region.**

AGI	Description	t-test p-value	Fold change
At2g21320	CONSTANS B-box zinc finger family protein	0.001073	37.45
At5g42220	ubiquitin family	0.01068	25.76
At4g00390	expressed protein	0.003535	14.37
At4g21090	ferredoxin family	0.047603	11.98
At4g16430	bHLH protein family	0.042494	5.765
At5g10480	protein phosphatase -related	0.014861	5.569
At5g51960	expressed protein	0.033587	5.407
At1g71770	polyadenylate-binding protein 5 (PABP5)	0.012837	4.989
At4g30660	stress responsive protein homolog	0.037245	4.24
At2g07170	expressed protein	0.004957	3.437

Table 5.2 The ten most differentially expressed significant genes up-regulated in the basal region of the globular stage embryo compared to the apical region.

AGI	Description	t-test p-value	Fold change
At1g04410	malate dehydrogenase, cytosolic	0.006428	6.711
At5g60960	pentatricopeptide (PPR) repeat-containing protein	0.006434	4.673
At4g30800	40S ribosomal protein S11 (RPS11B)	0.019381	4.098
At1g47260	transferase hexapeptide repeat family	0.000388	3.984
At3g57930	expressed protein	0.037292	3.289
At5g47930	ribosomal protein S27	0.040566	2.387
At1g06040	salt-tolerance protein	0.035337	2.381
At5g20890	chaperonin, putative	0.036937	2.278
At5g43750	expressed protein	0.04955	2.212
At5g08690	H+-transporting ATP synthase beta chain (mitochondrial)	0.046671	2.105

Further filtering of these genes was accomplished by calculating a fold-change between the expression values of the two regions for a particular gene. The 10 most up-regulated genes in the apical and basal regions passing the significance filter are shown in Tables 5.1 and 5.2. It is significant to note, even from this small sample of genes, that whilst the highest fold-change observed in the apical sample is ~37.5 times, in the basal sample the highest is for At1g04410, a cytosolic, malate dehydrogenase which is only 6.7 times more highly expressed in the basal sample compared to the apical sample. A range of functional groups are represented in the gene samples, although no one group is represented more highly than 2 out of the 10. No putative transcription factors are present in the up-regulated basal sample compared to 2 which are found in the apical sample; At2g21320, a CONSTANS B-box zinc finger family protein which is up-regulated by ~37.5 times, and At4g16430, a βHLH protein which is up-regulated ~5.8 times. Also of note in the apical sample are At5g42220, a ubiquitin family protein (with a predicted role in protein fate) up-regulated ~25.8 times, and At5g10480 (*PEPINO*), an *EMB* gene encoding a putative anti-phosphatase, which is up-regulated 5.6 times.

5.6.2 Heart cotyledon versus root comparison

As in the previous section, the heart stage cotyledon and root samples were normalised together using the root sample as the control. The resulting data were then filtered by significance using a Student's t-test with a maximum confidence level of 5% for genes whose expression was significantly different from a value of one. To correct for the occurrence of false positives a Benjamini and Hochberg false discovery rate, multiple testing corrections was used to adjust the p-values (Benjamini and Hochberg, 1995; GeneSpring version 7.2).

532 genes satisfied these criteria, and were sorted into those up-regulated and down-regulated in the cotyledon compared to the root (or *vice versa* for the root compared to the cotyledon).

Further filtering of these genes was accomplished by calculating a fold-change between the expression values of the two regions for a particular gene. The 10

Table 5.3 The ten most differentially expressed significant genes up-regulated in the cotyledon region of the heart stage embryo compared to the root region.

AGI	Description	t-test p-value	Fold change
At4g37750	ovule development protein aintegumenta (ANT)	0.001147	35.25
At5g56600	profilin 5	0.009519	3.128
At3g01910	sulfite oxidase	0.012128	2.851
At4g02840	small nuclear ribo protein Sm-D1-related protein	0.003713	2.748
At4g07950	DNA-directed RNA polymerase subunit -related	0.036269	2.606
At4g29840	threonine synthase (chloroplast)	0.000913	2.436
At5g47870	expressed protein	0.021925	2.134
At4g01370	mitogen-activated protein kinase (AtMPK4)	0.031034	2.117
At4g09550	expressed protein	0.038148	2.096
At5g10270	Cyclin-dependent kinase C;1 cdc2-like protein kinase	0.028007	1.924

Table 5.4 The ten most differentially expressed significant genes up-regulated in the root region of the heart stage embryo compared to the cotyledon region.

AGI	Description	t-test p-value	Fold change
At5g01870	lipid transfer protein, putative	0.043594	44.14
At1g48470	glutamine synthetase -related	0.017019	32.85
At5g55250	S-adenosyl-L-methionine:carboxyl methyltransferase family	0.003634	30.34
At3g03070	expressed protein	0.030517	20.36
At3g54260	expressed protein	0.035483	16.48
At1g73620	thaumatin-like protein (pathogenesis-related protein), putative	0.029497	12.52
At2g02510	expressed protein	0.04582	9.324
At1g32930	galactosyltransferase family	0.013409	8.287
At1g32790	RNA-binding protein, putative	0.041178	8.213
At3g63200	patatin-related	0.030779	8.17

most up-regulated genes in the cotyledon and root passing the significance filter are shown in Tables 5.3 and 5.4. The cotyledon gene sample is enriched with putative transcription factors (3 out of 10), however of these only At4g37750 (*AINTEGUMENTA*) which is ~35.25 times up-regulated in the cotyledon sample, displays a high fold-change, the others; At4g02840, a small nuclear ribo protein, is up-regulated ~2.75 times, and At4g07950, a DNA-directed RNA polymerase, is up-regulated only ~2.6 times. In marked contrast, the root gene sample all display notably high fold-changes. The root sample is enriched with metabolic genes, a group which includes the 3 highest fold-changes, from ~30 times up to the ~44 times higher in the root sample compared to the cotyledon, which is displayed by At5g01870, a putative lipid transfer protein. The root sample also includes 1 putative transcription factor, At1g32790, a putative RNA-binding protein which is up-regulated ~8.2 times in the root compared to the cotyledon.

5.6.3 Torpedo cotyledon versus root comparison

The torpedo-stage cotyledon and root samples were normalised together using the root sample as the control. The resulting data were then filtered by significance using a Student's t-test with a maximum confidence level of 5% for genes whose expression was significantly different from a value of one. To correct for the occurrence of false positives a Benjamini and Hochberg false discovery rate, multiple testing corrections was used to adjust the p-values (Benjamini and Hochberg, 1995; GeneSpring version 7.2).

1834 genes satisfied these criteria, and were sorted into those up-regulated and down-regulated in the cotyledon compared to the root (or *vice versa* for the root compared to the cotyledon).

Further filtering of these genes was accomplished by calculating a fold-change between the expression values of the two regions for a particular gene. The 50 most up-regulated genes in the cotyledon and root passing the significance filter are shown in Tables 5.5 and 5.6.

Table 5.5 The fifty most differentially expressed significant genes up-regulated in the cotyledon region of the torpedo stage embryo compared to the root region.

The torpedo-stage cotyledon and root samples were normalised together using the root sample as the control. The resulting data were then filtered by significance using a Student's t-test with a maximum confidence level of 5% for genes whose expression was significantly different from a value of one. To correct for the occurrence of false positives a Benjamini and Hochberg false discovery rate, multiple testing corrections was used to adjust the p-values (Benjamini and Hochberg, 1995; GeneSpring version 7.2).

1834 genes satisfied these criteria, and were sorted into those up-regulated and down-regulated in the cotyledon compared to the root.

Further filtering of these genes was accomplished by calculating a fold-change between the expression values of the two regions for a particular gene. The 50 most up-regulated genes in the cotyledon passing the significance filter are shown.

Table 5.5 The fifty most differentially expressed significant genes up-regulated in the cotyledon region of the torpedo stage embryo compared to the root region.

AGI	Description	t-test p-value	Fold change
At2g23170	expressed protein	0.000838	52.2
At5g48490	protease inhibitor/seed storage/lipid transfer protein (LTP) family	0.001415	31.42
At1g68780	leucine rich repeat protein family contains leucine rich-repeat (LRR) domains	0.002153	21.07
At2g03870	snRNP splicing factor -related	0.014837	16.72
At5g47500	pectinesterase family	0.013733	16.42
At4g37750	ovule development protein aintegumenta (ANT)	0.000381	13.41
At5g18790	ribosomal protein L33 - like	0.009354	12.32
At3g12203	serine carboxypeptidase -related	0.004209	10.77
At5g63530	copper chaperone (CCH)-related low similarity to copper homeostasis factor	0.005589	10.45
At1g08380	expressed protein	0.011549	9.174
At5g39860	bHLH protein putative DNA-binding protein	0.008136	9.09
At1g71695	peroxidase, putative	0.014996	8.428
At2g34430	photosystem II type I chlorophyll a /b binding protein	0.031424	8.273
At3g54560	histone H2A.	0.031955	7.663
At3g49670	leucine-rich repeat transmembrane protein kinase, putative CLAVATA1	0.017888	7.298

At1g78630	ribosomal protein L13p family	0.016467	7.175
At2g20820	auxin-regulated protein .	0.030591	7.046
At3g14190	expressed protein	0.040111	6.823
At5g04490	expressed protein	0.044855	6.586
At3g52730	ubiquinol--cytochrome-c reductase-related protein	0.029559	6.444
At5g66940	Dof zinc finger protein	0.000458	6.302
At5g61930	expressed protein	0.023844	6.179
At4g36870	BEL1-like homeobox 2 protein (BLH2)	0.04431	6.172
At2g02130	plant defensin protein, putative	0.016127	6.093
At1g78820	curculin-like (mannose-binding) lectin family	0.026986	6.018
At3g14110	tetratricopeptide repeat (TPR)-containing protein	0.030088	6.002
At5g64530	No apical meristem (NAM) protein family	0.001633	5.799
At3g51660	Macrophage migration inhibitory factor (MIF) family	0.011083	5.631
At3g58180	expressed protein	0.018644	5.443
At4g15620	expressed protein	0.041922	5.38
At4g10100	auxin-regulated protein	0.039056	5.362
At5g54630	expressed protein	0.008897	5.34
At4g12800	probable photosystem I chain XI precursor	0.027934	5.237
At2g06520	expressed protein	0.003584	5.229
At3g24570	peroxisomal membrane protein family	0.018189	5.151

At1g10060	branched-chain amino acid aminotransferase -related	0.008357	5.102
At1g29070	plastid ribosomal protein L34 precursor -related	0.03356	5.06
At3g03920	glycine-rich protein	0.048425	5.054
At4g29735	expressed protein	0.018419	4.681
At1g55670	photosystem I subunit V precursor -related	0.031079	4.627
At1g35680	50S ribosomal protein L21 chloroplast precursor (CL21)	0.001121	4.615
At3g45770	oxidoreductase, zinc-binding dehydrogenase family	0.000833	4.576
At2g26330	receptor-related protein kinase, ERECTA	0.021804	4.523
At4g24150	hypothetical protein	0.027552	4.491
At5g39790	5'-AMP-activated protein kinase, beta-1 subunit	0.032003	4.461
At3g22630	20S proteasome beta subunit D (PBD1)	0.022456	4.402
At2g33800	ribosomal protein S5 family	0.020719	4.371
At4g34620	ribosomal protein S16p family	0.0000443	4.347
At4g09830	expressed protein	0.033349	4.226
At1g51650	epsilon subunit of mitochondrial F1-ATPase	0.044062	4.206

Table 5.6 The fifty most differentially expressed significant genes up-regulated in the root region of the torpedo stage embryo compared to the cotyledon region.

The torpedo-stage cotyledon and root samples were normalised together using the root sample as the control. The resulting data were then filtered by significance using a Student's t-test with a maximum confidence level of 5% for genes whose expression was significantly different from a value of one. To correct for the occurrence of false positives a Benjamini and Hochberg false discovery rate, multiple testing corrections was used to adjust the p-values (Benjamini and Hochberg, 1995; GeneSpring version 7.2).

1834 genes satisfied these criteria, and were sorted into those up-regulated and down-regulated in the root compared to the cotyledon.

Further filtering of these genes was accomplished by calculating a fold-change between the expression values of the two regions for a particular gene. The 50 most up-regulated genes in the root passing the significance filter are shown.

Table 5.6 The fifty most differentially expressed significant genes up-regulated in the root region of the torpedo stage embryo compared to the cotyledon region.

AGI	Description	t-test p-value	Fold change
At3g63040	expressed protein	0.001258	8.737
At2g23510	transferase family	0.00000447	7.686
At3g13520	arabinogalactan-protein (AGP12)	0.017381	7.368
At1g20450	dehydrin (ERD10)	0.003299	6.708
At3g54260	expressed protein	0.000325	5.995
At4g30850	expressed protein	0.025604	5.61
At5g14920	expressed protein	0.012758	4.915
At1g58270	expressed protein	0.001044	4.687
At3g48460	GDSL-motif lipase/hydrolase protein similar to lipase	0.018973	4.555
At1g74500	bHLH protein family	0.000402	3.949
At5g35630	glutamate-ammonia ligase precursor (chloroplast)	0.00427	3.654
At5g17430	ovule development protein, putative similar to aintegumenta	0.005328	3.628
At5g35940	jacalin lectin family	0.0000213	3.626
At3g53450	hypothetical protein lysine decarboxylase	0.02311	3.55
At1g75710	expressed protein	0.018169	3.419

At4g10640	calmodulin-binding protein family	0.0000309	3.246
At4g32870	expressed protein	0.038469	3.219
At1g25530	Transmembrane lysine and histidine specific amino acid transporter protein	0.001659	3.198
At5g11930	glutaredoxin protein family	0.001676	3.144
At5g18310	expressed protein	0.03409	3.113
At4g13030	expressed protein	0.033043	3.083
At4g14130	xyloglucan endotransglycosylase (XTR7)	0.000114	3.034
At2g46410	myb-related protein CAPRICE (CPC)	8.36E-05	3.032
At3g63200	patatin-related	0.003522	2.953
At2g01770	membrane protein -related	0.000454	2.952
At2g17640	acetyltransferase	0.037655	2.877
At3g03340	expressed protein	0.040951	2.872
At1g31950	sesquiterpene synthase/cyclase family	0.004715	2.856
At1g64640	plastocyanin-like domain containing protein	0.015199	2.808
At3g50870	GATA zinc finger protein	0.0258	2.796
At5g55250	S-adenosyl-L-methionine:carboxyl methyltransferase family	0.042493	2.754
At1g52690	late embryogenesis abundant (LEA) protein	0.00777	2.701
At1g67530	armadillo repeat containing protein	0.003924	2.7
At2g45510	cytochrome p450, putative	0.017945	2.628
At2g25980	jacalin lectin family	0.0365	2.626

At1g79840	homeodomain protein GLABRA2 (GL2)	0.019134	2.616
At1g35290	thioesterase-related protein	0.00223	2.611
At3g15240	expressed protein	0.000260	2.584
At1g78580	UDP-forming (trehalose-6-phosphate synthase)	0.011816	2.564
At2g47780	expressed protein	0.001278	2.528
At2g26520	expressed protein	0.030302	2.498
At3g16857	response regulator protein family; similar to ARR1 protein	0.001966	2.488
At2g19590	1-aminocyclopropane-1-carboxylate oxidase (ACC oxidase)	0.00359	2.464
At5g41620	expressed protein	0.023059	2.444
At2g17840	senescence-associated protein 12 -related	0.002181	2.442
At1g02810	pectinesterase family	0.009217	2.44
At4g34760	auxin-induced (indole-3-acetic acid induced) protein family	0.00118	2.335
At1g28400	expressed protein	0.0000314	2.322
At1g49390	oxidoreductase, 2OG-Fe(II) oxygenase family	0.046114	2.3
At3g13175	expressed protein	0.008691	2.242

Figure 5.11 Validation of GeneSpring analysis by RT-PCR analysis of aRNA.

RT-PCR analysis on aRNA from torpedo stage embryo's after three rounds of amplification to validate the GeneSpring analysis to uncover significant genes, which exhibit differential expression between the cotyledon region (Co) and the root region (Rt). *ACT3* a predicted constitutive gene was used as a control.

Apical genes

The five most differentially expressed significant genes up-regulated in the cotyledon region of the torpedo stage embryo compared to the root region.

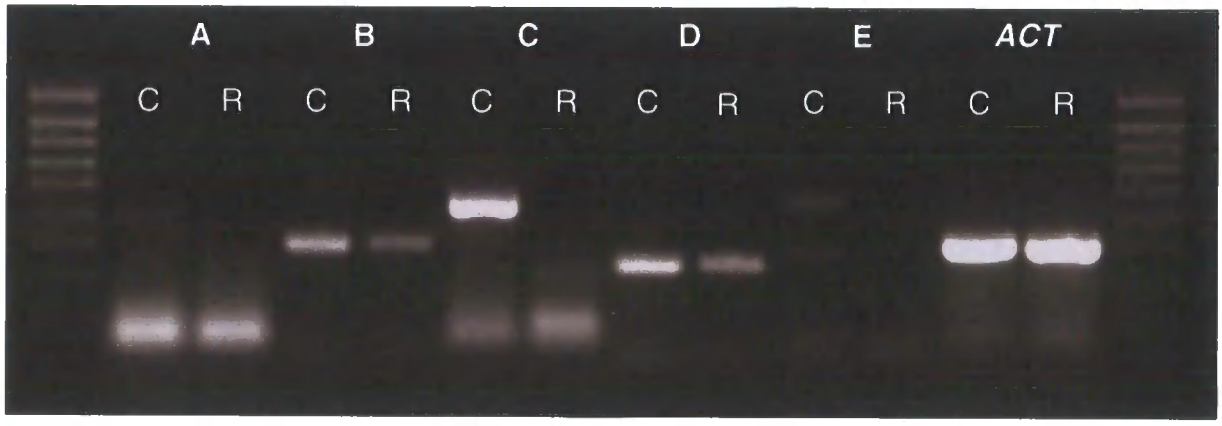
A – At2g23170 (expressed protein)
B - At5g48490 (protease inhibitor/seed storage/lipid transfer protein)
C - At1g68780 (leucine rich repeat protein family)
D - At2g03870 (snRNP splicing factor –related)
E - At5g47500 (pectinesterase family)
ACT3 - Control

Basal genes

The five most differentially expressed significant genes up-regulated in the root region of the torpedo stage embryo compared to the cotyledon region.

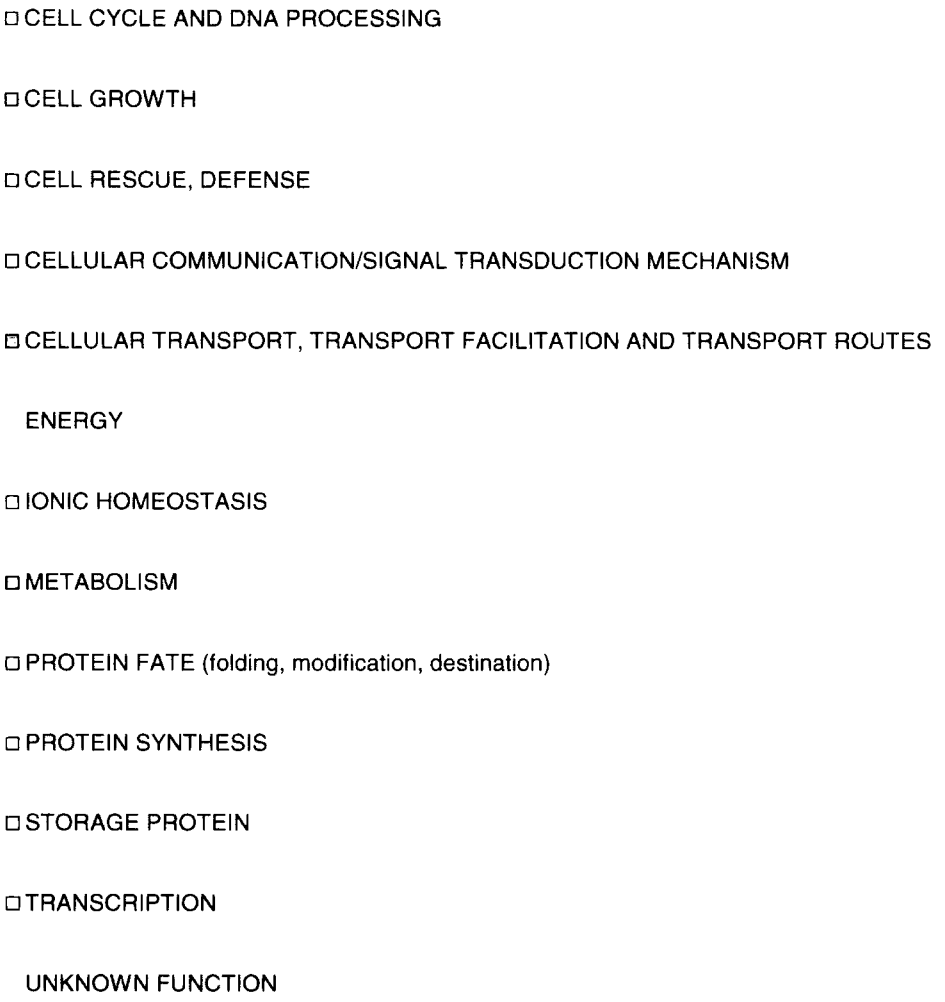
F - At3g63040 (expressed protein)
G - At2g23510 (transferase family)
H - At3g13520 (arabinogalactan-protein (AGP12))
I - At1g20450 (dehydrin (ERD10))
J - At3g54260 (expressed protein)
ACT3 - Control

(i) Apical

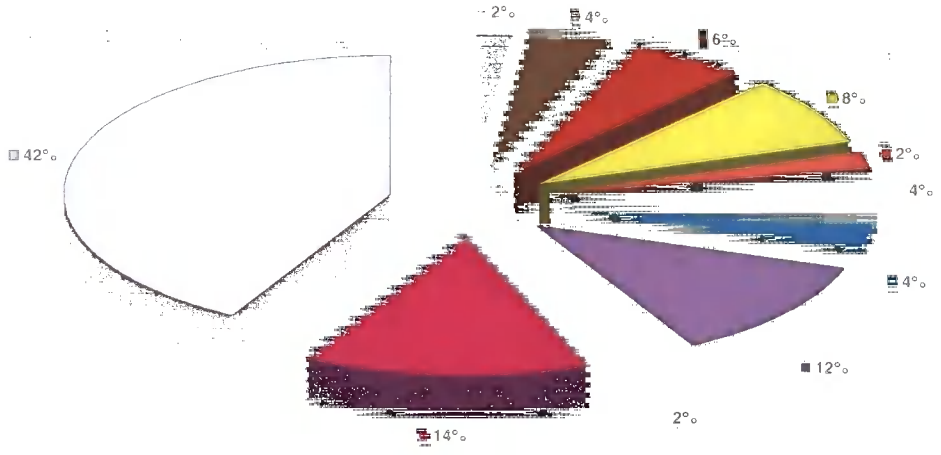


(ii) Basal

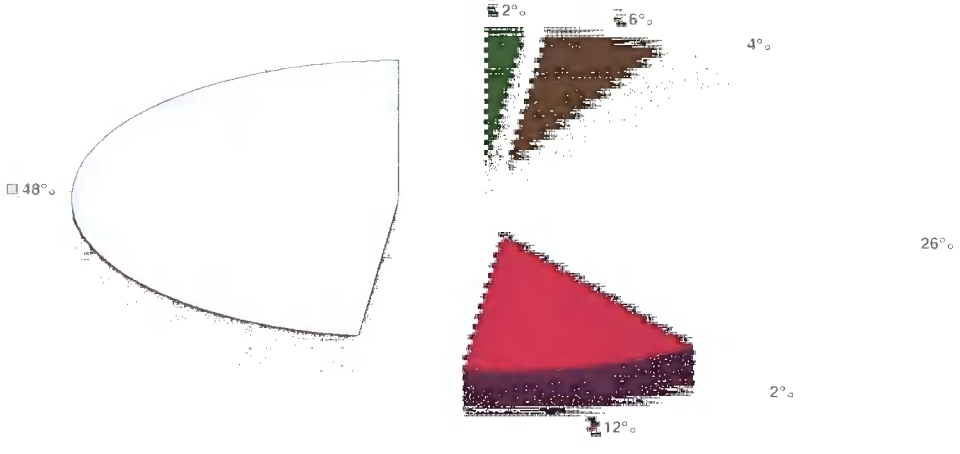
Figure 5.12 Functional annotation of the 50 most differentially expressed (by fold-change) significant genes ($p<0.05$) between the cotyledon and root regions of the torpedo stage embryo.



Functional classification of the 50 most up-regulated significant genes ($p<0.05$) in the cotyledon region of the torpedo stage embryo compared to the root region.



Functional classification of the 50 most up-regulated significant genes ($p<0.05$) in the root region of the torpedo stage embryo compared to the cotyledon region.



In order to validate these findings for the torpedo cotyledon versus root comparison, RT-PCR analysis was performed on aRNA after three rounds of amplification (as described in Section 2.10.3), on the 5 most highly up-regulated genes in both the cotyledon and root samples. The results of this analysis are shown in Figure 5.11. As can be seen the RT-PCR analysis confirms the expression patterns highlighted by the GeneSpring analysis. Of the 5 genes predicted to be up-regulated in the cotyledon, 3 produced products exclusively from the cotyledon sample, the other 2 produced a product for both the cotyledon and the root sample, although the amount of product was greater for the cotyledon sample. Of the 5 genes predicted to be up-regulated in the root, 3 produced products exclusively from the root sample, the other 2 failed to produce products, although test reactions showed this to be a result of inadequate primers (results not shown). All reactions used *ACT3* as a control, which was present constitutively in both samples.

The 50 most up-regulated genes in the cotyledon and root passing the significance filter were assigned functional annotation using information from http://mips.gsf.de/proj/thal/db/. Figure 5.12 displays these functional classifications of the up-regulated genes between the tissue samples. In general the cotyledon sample has a much larger range of functional groups represented compared to the root sample. The most enriched groups in the cotyledon sample are transcription (14%) and protein synthesis (12%), additional groups enriched at a lower level include energy (8%) and signal transduction (6%). In contrast only 2 main groups are enriched in the root sample these being metabolism (26%) and transcription (12%). Cell rescue and defence response genes also comprise 6% of the root sample and 4% of the cotyledon sample. In both samples the largest group is that of unknown function which comprises 42% of the cotyledon sample and 48% of the root sample.

5.6.4 Globular apical versus Torpedo SAM

In addition to comparisons between the apical and basal regions of the different stages of embryogenesis, it was also decided to compare the apical region of the globular stage embryo to the shoot apical meristem region of the torpedo

stage embryo. The reasoning behind this comparison is that in addition to being the region from which the cotyledons are eventually derived, it has also been shown that, despite no morphologically recognisable structure, expression of essential SAM genes is nevertheless centred in this region (Barton and Poethig, 1993; Long *et al.*, 1996; Mayer *et al.*, 1998). The clonal destiny of this region to form the shoot apical meristem is also predicted to be determined in the early globular stage, therefore making this a valid comparison between a presumptive SAM at the globular stage and its more developed form at the torpedo stage (Christianson, 1986; Poethig *et al.*, 1986).

The globular-stage apical region and the torpedo-stage shoot apical meristem samples were normalised together using the globular stage apical sample as the control. The resulting data was then filtered by significance using a Student's t-test with a maximum confidence level of 5% for genes whose expression was significantly different from a value of one. To correct for the occurrence of false positives a Benjamini and Hochberg false discovery rate, multiple testing corrections was used to adjust the p-values (Benjamini and Hochberg, 1995; GeneSpring version 7.2).

921 genes satisfied these criteria, and were sorted into those up-regulated and down-regulated in the torpedo stage shoot apical meristem compared to the globular stage apical region (or *vice versa* for the globular apical region compared to the torpedo shoot apical meristem).

Further filtering of these genes was accomplished by calculating a fold change between the expression values of the two regions for a particular gene. The 10 most up-regulated genes in the torpedo stage shoot apical meristem and the globular stage apical region, passing the significance filter, are shown in Tables 5.7 and 5.8. From these samples it is difficult to isolate potentially important meristematic genes particularly as the majority of the globular-stage apical sample is of unknown function. In the torpedo-stage meristem sample the characterised meristematic transcription factor At1g622360 (*STM*) is up-regulated 21.44 times. Interestingly the torpedo-stage meristem sample is enriched (2 out of 10) with known cytoskeleton genes; At1g04820 (*TUA4*), a tubulin alpha-2/alpha-4 chain, and At2g35630 (*MOR1*), microtubule

Table 5.7 The ten most differentially expressed significant genes up-regulated in the meristem of the torpedo stage embryo compared to the apical region of the globular stage embryo.

AGI	Description	t-test p-value	Fold change
At1g62360	*SHOOT MERISTEMLESS (STM)*	0.000477	21.44
At5g38195	protease inhibitor/seed storage/lipid transfer protein (LTP) family	0.020827	20.06
At5g03350	expressed protein	0.046612	19.75
At5g18620	DNA-dependent ATPase	0.001223	13.39
At2g35900	expressed protein	0.018532	8.33
At1g69220	serine/threonine protein kinase	0.006282	5.777
At1g04820	tubulin alpha-2/alpha-4 chain (TUA4)	0.023085	5.556
At2g35630	microtubule organization 1 protein (MOR1)	0.023514	5.34
At3g27190	uracil phosphoribosyltransferase -related	0.015652	5.099
At3g63490	ribosomal protein L1p family	0.005844	4.816

Table 5.8 The ten most differentially expressed significant genes up-regulated in the apical region of the globular stage embryo compared to the meristem of the torpedo stage embryo.

AGI	Description	t-test p-value	Fold change
At2g30810	gibberellin-regulated protein -related	0.001391	84.73
At4g29285	expressed protein	0.007314	69.44
At2g38900	protease inhibitor -related	0.006098	68.34
At5g06190	expressed protein	0.010398	65.1
At4g09030	arabinogalactan-protein (AGP10)	0.01171	63.55
At3g24510	hypothetical protein	0.006217	53.28
At5g08040	expressed protein	0.026082	51.41
At2g45110	beta-expansin, putative	0.014973	50.64
At3g46560	small zinc finger-related protein TIM9	0.029884	50.05
At5g05370	ubiquinol-cytochrome C reductase complex ubiquinone-binding protein	0.015141	47.95

organisation protein 1, both of which are up-regulated more than 5 times in the torpedo-stage, meristem sample compared to the globular-stage, apical sample.

5.6.5 7 dpg Seedling cotyledon versus root comparison

The seedling cotyledon and root samples were normalised together using the root sample as the control. The resulting data were then filtered by significance using a Student's t-test with a maximum confidence level of 5% for genes whose expression was significantly different from a value of one. To correct for the occurrence of false positives a Benjamini and Hochberg false discovery rate, multiple testing corrections was used to adjust the p-values (Benjamini and Hochberg, 1995; GeneSpring version 7.2).

13,357 genes satisfied these criteria, and were sorted into those up-regulated and down-regulated in the cotyledon compared to the root (or *vice versa* for the root compared to the cotyledon)

Further filtering of these genes was accomplished by calculating a fold-change between the expression values of the two regions for a particular gene. The 10 most up-regulated genes in the cotyledon and root passing the significance filter are shown in Tables 5.9 and 5.10. The cotyledon and root samples show very clear trends even at this sample size. In the cotyledon three major functional classes are represented, namely; chloroplastic metabolism/biosynthesis, cell rescue and defence response, and energy. In the root sample only two functional classes are represented these are cell rescue and defence response, and metabolism. All the genes in these samples are backed by very high fold-changes.

5.6.6 Analysis of significant genes across developmental stages

The apical-basal comparisons made at each developmental stage, shown in the preceding sections highlighted a significant number of genes with potential

Table 5.9 The ten most differentially expressed significant genes up-regulated in the cotyledon of a 7dpg seedling compared to the root.

AGI	Description	t-test p-value	Fold change
At3g45140	lipoxygenase (LOX2)	0.000133	315.9498
At5g13630	cobalamin biosynthesis protein	0.0000607	289.384
At3g25760	allene oxide cyclase family; dehydration (ERD12) protein	0.000118	281.4639
At1g23130	Pathogenesis-related protein Bet v I family	0.0000252	263.3356
At3g01500	carbonic anhydrase, chloroplast precursor	0.000112	240.1065
At1g42970	glyceraldehyde-3-phosphate dehydrogenase	0.000502	233.6202
At1g68010	NADH-dependent hydroxypyruvate reductase	0.0000684	228.7297
At3g62410	CP12 protein precursor-related protein (chloroplast)	0.000147	210.8997
At1g31580	ORF1	0.000210	209.519
At4g03280	Rieske FeS protein (component of cytochrome B6-F complex)	0.000113	204.6475

Table 5.10 The ten most differentially expressed significant genes up-regulated in the root of a 7dpg seedling compared to the cotyledon.

AGI	Description	t-test p-value	Fold change
At2g38390	peroxidase, putative	0.0000250	99.86
At5g05960	protease inhibitor/seed storage/lipid transfer protein (LTP) family	0.000467	97.409
At5g63600	flavonol synthase, putative, oxygenase superfamily	0.000462	96.17
At2g01530	major latex protein (MLP)-related	0.000961	94.886
At2g44790	uclacyanin II	0.000494	94.742
At1g49860	glutathione transferase	0.000994	93.703
At2g05510	glycine-rich protein	0.001922	90.514
At4g11190	disease resistance response protein family/ dirigent protein family	0.000471	89.558
At5g01870	lipid transfer protein	0.001422	81.853
At1g28290	prolin-rich protein -related	0.002808	81.123

Table 5.11 A comparison of differentially expressed genes, between tissues and developmental stages.

A comparative analysis was performed between the significant ($p<0.05$) apical and basal gene sets at different developmental stages to assess the degree of overlap. To determine whether these overlaps reflected a distinct population of purely apical or basal genes and not random chance a Chi^2 test was performed.

Pale yellow shading suggests that no more genes are in common than would be expected if the two treatments were independent.

Pale green shading suggests that more genes are in common than would be expected if the two treatments were independent (i.e. a significant overlap).

Blue shading highlights the total number of significant genes present in a specific tissue type at a given developmental stage.

Table 5.11. A comparison of differentially expressed genes, between tissues and developmental stages.

		Globular		Heart		Torpedo			Seedling	
		Apical	Basal	Apical	Basal	Apical	Basal	Meristem	Apical	Basal
Seedling	Basal	81	101	124	83	334	285	83	0	7778
	Apical	69	84	74	38	371	119	64	5579	
Torpedo	Meristem	1	6	2	2	11	9	228		
	Basal	5	6	10	9	0	584			
	Apical	25	14	19	13	1250				
Heart	Basal	1	1	0	187					
	Apical	5	5	345						
Globular	Basal	0	305							
	Apical	280								

importance in apical-basal polarity. It can be hypothesised that whilst some genes will be significant for limited periods of developmental time there will also be a sub-set of genes which have a spatial localisation for an extended period of developmental time. Therefore a comparative analysis was performed between the apical and basal gene sets at different developmental stages to assess the degree of overlap. To determine whether these overlaps reflected a distinct population and not random chance a Chi^2 test was performed. This analysis is summarised in Table 5.11. A number of the overlaps were very close to the 95% confidence limit set but it was decided to maintain this as a standard. The results are inconclusive as to the existence of distinct apical and basal gene sets through development, indeed, the only overlap of significant genes was between the torpedo-stage and the seedling, and this was not confined to solely an apical and basal overlap.

5.7 Discussion

Microarrays are a very powerful technique allowing the expression profiles of thousands of genes to be monitored simultaneously. The amount of data obtainable from a microarray experiment is so extensive as to make a general summary difficult. Indeed the major difficulty associated with microarray technology is not the generation of data but the ability to differentiate between significant biological changes in gene expression and experimental error. The extraction of significant and valid information from microarray data is a product of two major phases in the processing of the raw data; data transformation and normalisation, and data analysis.

The aims of this work did not extend to an assessment of different data transformation approaches, and although it is a field in which a great deal of research emphasis is placed, as yet no gold-standard approach has emerged. However, some form of normalisation is required to ensure that any differential expression observed is not the result of 'obscuring variation' from printing, hybridisation or scanning artefacts (Irizarry *et al.*, 2003b).

Conceptually normalisation can be regarded as a similar process to the adjustment of expression levels relative to that of a control gene(s) assumed to be constant between samples, as is the case for quantitative reverse transcription PCR (RT-PCR) (Quackenbush, 2002; Brunner *et al.*, 2004).

A major difference between the default Affymetrix MAS 5.0 normalisation used in Chapter 4 and the RMA normalisation used in this Chapter is that of background correction. Prior to normalisation, the MAS 5 algorithm uses the ideal mismatch system of background correction (Affymetrix, 2002). This uses the MM probes to correct the PM probes for background noise, the ideal mismatch procedure takes into account that approximately 1/3 of MM are greater than the PM due to target sensitivity (Affymetrix, 2002; Irizarry 2003b). Comparison of background corrections found that while normalisation reduced the general variability, background correction added to the variability, with the ideal mismatch procedure increasing the variability the most (Bolstad, 2002; Irizarry *et al.*, 2003b). RMA normalisation therefore uses only the PM probes, sacrificing some degree of accuracy (attenuation of fold-changes at low abundances) for large gains in precision (Irizarry *et al.*, 2003b; Wu *et al.*, 2004). Recently a modified version of RMA has been developed that creates a background estimate (pseudo-MM) based on a model using the G-C content of the probe, thus retaining the precision of RMA but increasing the accuracy at low abundance levels (Wu *et al.*, 2004). GC-RMA represents a significant improvement over RMA and would be the preferred system for future analysis.

Comparisons using spike-in and dilution data sets have been used to compare the RMA and MAS 5.0 normalisation methods. In general RMA was seen to add a slight bias to signal estimates, which is overcome by a significant reduction in variance. RMA compresses log-fold changes by 10-20% compared to MAS 5.0, but the same estimates show a 9 times larger variability in the MAS 5.0 normalised data. Compared to MAS 5.0 normalisation the advantages RMA normalisation offers for the global analysis presented in this chapter include a 5-fold reduction of within replicate variance, more consistent estimates of fold-change and increased sensitivity in detecting differential expression from fold-changes (Irizarry *et al.*, 2003a; Irizarry *et al.*, 2003b; Cope *et al.*, 2004).

Having used RMA normalisation to differentiate between biological variation and experimental error, allowing comparison to be made between the arrays used in this study, the next phase is that of data analysis.

5.7.1 Relatedness of embryonic developmental stages

In order to test the hypothesis that each developmental stage under investigation has a distinct transcriptional profile a condition tree clustering analysis was performed using Spearman correlation to rank the similarity (Section 5.3).

All tissue types sampled from the same developmental stage clustered together as opposed to all the apical regions clustering separately from the basal regions. This demonstrates that in terms of overall transcriptional profile there appears to be a greater input from the temporal expression patterns than the spatial expression patterns. This data fits with an established model based on RNA hybridization studies in tobacco, which suggests that while there are significant populations of organ-specific transcripts, 60-77% of plant genes are present in heterologous organs (Girke *et al.*, 2000). Goldberg *et al.* (1989) demonstrated that distinct mRNA sets are temporally regulated during embryogenesis, with expression restricted to specific developmental stages.

Ma *et al.* (2005) conducted a study into the transcriptional profiles of specific organs, removing the temporal dimension by focusing on the mature *Arabidopsis* plant. Relatedness was shown to correlate with the developmental relationship of the organs, with roots and leaves clustering separately in distinct groups, thus demonstrating the presence of distinct spatial patterns of gene expression. While all the organs analysed can be considered to represent a single developmental stage (i.e. mature), a transcriptional profile of the germinating seed was also conducted and shown to cluster in a distinct group on its own.

Beemster *et al.* (2005) present a study of the relatedness of leaf samples over a developmental time-course. Tissue samples from different aged leaves

clustered into three distinct groups, representing the developmental phases: proliferation, expansion and mature, as revealed by kinematic analysis. This demonstrates the presence of distinct temporal patterns of gene expression in a specific developing organ.

Scheidl *et al.* (2002) suggested that any difference in expression profile observed between amplified and non-amplified samples was the consequence of a global reduction in transcript length resulting from priming with random hexamers. It might therefore be expected that the embryonic samples, which underwent amplification, would cluster separately from the seedling samples. However, the condition tree does not show this distinction; instead the torpedo-stage samples cluster with the seedling samples. Given that all the embryonic samples underwent an identical amplification procedure, this result suggests a distinction in transcriptional profile between early and mid/late embryogenesis rather than any technical bias.

The globular- and heart-stage samples represent the early stages of embryogenesis characterised by pattern formation and morphogenesis, with the axes of the plant body plan defined and the organ systems formed. In contrast, the torpedo-stage represents mid/late embryogenesis characterised by maturation and the accumulation of storage reserves in preparation for developmental arrest (Lindsey and Topping, 1993). Given that the plant body pattern is already established and that the seedling continues to undergo a maturation process, it would be expected to have a transcriptional profile more similar to the torpedo-stage than the early embryonic stages, as confirmed by the condition tree analysis.

5.7.2 Transcriptional changes along an embryonic developmental time-course

Having established that the arrays clustered according to a temporal rather than spatial dominance of transcriptional profile, it was decided to analyse the transcription changes in a specific region along a developmental time-course in a similar way to that achieved by Beemster *et al.* (2005).

The normalised data for the apical and basal regions were plotted on a log scale against developmental stage as shown in Figures 5.3 and 5.7 respectively. These figures include data for all the genes present on the GeneChip®, with no filtering imposed as to expression cut-off threshold or statistical significance. Both time-course figures show the majority of genes to be expressed at the same level throughout the developmental stages centred on the default expression value of 1. However each developmental stage is represented by a distinct outline of overall up- and down-regulation in both the apical and basal region, suggesting the strong temporal profile highlighted by the condition tree analysis.

Functional annotation of genes, which show up-regulation along the developmental time-course, highlighted some interesting general trends in the data. However, a measure of caution must be applied to these observations for the following reasons: statistical significance and confidence of annotation. Firstly, in order not to be prohibitively restrictive the significance filter was relaxed with a 95% confidence required at only one of the three developmental stages. This allows a general overview of the time-course but has the potential to allow non-statistically significant results. Secondly, only approximately 10% of genes in the *Arabidopsis* genome (~2500) have had their function deduced or confirmed by direct experimental analysis, therefore the vast majority of functions are putative and assigned on the basis of sequence similarity (Hilson, 2003). A powerful example of how functional annotation based on sequence similarity must be regarded with caution, is presented by Martinoia *et al.* (2002). The ABC-transporter superfamily contains more than 100 members in *Arabidopsis*, with strong sequence similarity in many cases suggestive of functional redundancy. However, recent research has shown that family members participate in a wide range of non-related functions, contrary to sequence similarity, which might have suggested a family of functionally redundant proteins involved in detoxification processes (Henikoff *et al.*, 1997; Arabidopsis Genome Initiative, 2000; Martinoia *et al.*, 2002)

Along the apical time-course, genes with a functional role in energy, predominantly those involved in photosynthesis and carbon fixation, were up-

regulated at every stage. This is in accordance with *in situ* hybridisation studies that show an increasing abundance of chloroplastic gene transcripts progressively from the proembryo stage of embryogenesis, with the highest concentration observed in the cotyledons of the mature embryo (Degenhardt *et al.*, 1991). The peak in transcript abundance in the mature embryo corresponds with the fate of the cotyledons as the initial photosynthetic organs of the seedling, and allows photosynthesis to commence promptly post-germination (Degenhardt *et al.*, 1991; Raghavan, 1997). Chloroplastic gene transcripts are also present as up-regulated in the basal time-course, which corresponds with the constitutive pattern of expression observed in developing embryos of *Gossypium hirsutum* (Borroto and Dure, 1986), *Glycine max* (Chang and Walling, 1991, 1992) and *Arabidopsis thaliana* (Degenhardt *et al.*, 1991).

Another significant change in the apical time-course is the up-regulation of genes involved in protein synthesis, between the heart-stage and the torpedo-stage, this could be representative of the transition from early embryogenesis to the maturation and protein accumulation characteristic of mid/late embryogenesis (Lindsey and Topping, 1993).

Along the basal time-course, the transition from heart- to torpedo-stage is accompanied by the significant up-regulation of genes encoding proteins involved in cell growth, specifically the growth of cell walls. This group included a substantial number of hydroxyproline-rich glycoproteins (HRGPs), including expansins, which are regarded as key regulators of wall extension, causing 'wall loosening' thus allowing turgor-driven expansion (Cosgrove, 2000; Showalter, 2001; Li *et al.*, 2003). The up-regulation of this group of genes may reflect the elongation, which the embryo undergoes between the heart-stage and the torpedo-stage of embryogenesis (Leyser and Day, 2003). Expansins have also been implicated in desiccation tolerance and the up-regulation at the torpedo-stage of embryogenesis may also represent a stage in the preparation for developmental arrest and desiccation (Jones and McQueen-Mason, 2004).

In 1960, Jacob and Monod described the bacterial *lac* operon, a cluster of genes showing highly co-ordinated gene expression and functional interaction (Jacob and Monod, 1961). While operons are not generally found in eukaryotes,

the advent of global gene expression analysis has revealed that the highly co-ordinated expression of genes with related and interacting functions also occurs in eukaryotes (Niehrs and Pollet, 1999).

Cluster analysis is a potentially powerful technique with which to analyse microarray data. It has been utilised to identify groups of genes with similar expression patterns and by extension infer biological function to unknown genes through association with genes of known function (Eisen *et al.*, 1998; Hughes *et al.*, 2000; Wu *et al.*, 2002). Recently, Ren *et al.* (2005) have suggested that 5-10% of the *Arabidopsis* genome is composed of highly co-expressed, physically adjacent genes, the majority of which are not predicted to have the same function, emphasising that functionally 'guilt by association' must be treated with caution.

K-means clustering was performed on the filtered apical and basal time-course gene sets, in both cases 7 distinct dynamic expression patterns were observed along the time course. The clustering program used assigned genes to clusters based on a user defined cluster number, and did not create these clusters if the number was defined as 7. Hennig *et al.* (2004) predicted 9 models of dynamic expression pattern for genes involved in reproductive development of *Arabidopsis,* and used an alternative clustering package to assign their selected genes to the model class of best fit. This would appear to be an improvement in terms of selecting genes, as it would not create multiple clusters of genes showing essentially similar expression profiles. The number of time-points available is also of critical importance in cluster analysis; the 3 included here are the minimum and, as can be seen, produces clusters with considerable associated noise. Beemster *et al.* (2005) conducted cluster analysis on significantly modulated genes during leaf development from 9 to 31 days after sowing. Ten time-points were utilised and 16 very well defined clusters of 20 or more genes were produced. The acquisition and inclusion of a further time-point for the cotyledonary-stage of development would be predicted to greatly enhance the fidelity of clustering.

Individual clusters were probed with lists of transcription factor family members, however, no particular over-representation was observed in any of the clusters,

which could be a consequence of the amount of noise in the clusters masking such detail. Co-expression (genes with similar expression patterns) is substantially different from co-regulation (genes that are regulated by a common transcription factor) a distinction sometimes obscured by 'guilt through association' studies. Yeung *et al.* (2004), conducted model based clustering analysis on yeast, and concluded that between 50 and 100 microarray experiments are required to accurately identify co-regulated genes, and furthermore, predict that the correspondence between co-expression and co-regulation would be lower in more complex organisms.

5.7.3 Apical-basal tissue comparisons

Microarray data provide a powerful resource with which to uncover genes of potential importance on which to conduct further research. The control mechanisms underlying apical-basal polarity are a central theme in this work. Previously a number of putative transcription factors were isolated from the data and have been analysed by promoter::GUS fusion. It was decided to use GeneSpring to uncover statistically significant genes, which show differential expression between the apical (cotyledon) and basal (root) samples along the developmental time-course as a resource for future analysis. An additional aim was to deduce whether some of these apical and basal genes represented an organ specific gene-set throughout embryogenesis.

Casson *et al.* (2005) presented an analysis of spatially expressed putative transcription factors at the globular and heart-stages of embryogenesis. They isolated a gene (At2g21320) encoding a CONSTANS-like B-box zinc finger protein, expressed predominantly in the apical region. This gene emerged from the GeneSpring analysis of the Casson *et al.* (2005) data as the most differentially expressed significant gene in the globular apical region, providing a level of validation to the analysis. A number of other putative transcription factors emerged from this analysis, which could be investigated further.

As a non-embryonic comparison, data from 7 dpg seedling cotyledons and roots produced by the AtGenExpress Consortium, and provided by

NASCArrays at http://www.affymetrix.arabidopsis.info, were analysed. Compared to the embryonic samples, the seedling analysis produced extremely high fold-changes. No direct analysis was performed between embryonic data and seedling data so this observation would not be due to any technical effect of amplification or a dose-dependent effect. It is possibly suggestive that spatial gene expression is more pronounced in the seedling than the embryo. However as less embryonic genes passed the significance filters, it is possible that more variability exists between arrays and thus some of the genes exhibiting high-fold changes are not deemed statistically significant.

A comparison was also made between the apical region of the globular-stage embryo and the shoot apical meristem region of the torpedo stage embryo. Expression of essential SAM genes has been detected in the apical region of the globular-stage embryo and its clonal destiny to form the shoot apical meristem is also predicted to have been determined, therefore making this a valid comparison between a presumptive SAM at the globular stage and its more developed form at the torpedo stage (Christianson, 1986; Poethig *et al.*, 1986). The presence of *SHOOT MERISTEMLESS (STM)* in the up-regulated genes of the SAM indicates that the shoot apical meristem was successfully captured. Caution is required when analysing this data, as the initial laser capture microdissection step was not precise enough to achieve specific capture of the SAM and therefore some degree of contamination from the surrounding tissue is expected. Interestingly two known cytoskeletal genes were up-regulated in the SAM compared to the globular-stage apical region, namely *α-TUBULIN4* (*TUA4*; Kopczak *et al.*, 1992) and *MICROTUBULE ORGANIZATION1* (*MOR1;* Whittington *et al.*, 2001). Looking at the Affymetrix MAS 5.0 mean signal values it would appear that *TUA4* is constitutively expressed throughout the torpedo-stage embryo (mean signal values between ~193 and ~295) but is probably not expressed in the globular-stage apical region (mean signal value ~24), thus accounting for the high fold-change observed. *MOR1* appears to be expressed in the globular-stage apical region (mean signal value 119) and at a similar level in the torpedo-stage cotyledons and root (mean signal values ~142 and ~151), but appears much more abundant in the SAM (mean signal value ~802). The expression across tissues and developmental stages is consistent with Whittington *et al.* (2001), who

conclude that *MOR1* is required at all stages of development and in all organs. The 3 replicates performed for the torpedo SAM are very consistent in their signal values for *MOR1* suggesting that this could be an interesting result but further investigation is required.

The analysis of up-regulated genes in my torpedo-stage embryo cotyledons and root data provides a good basis for future work on this project. As an initial validation of the expression patterns seen in the microarray data I carried out RT-PCR on the top 5 differentially expressed cotyledon and root genes. Apart from 2 root genes for which more suitable primers are required to be designed, all the expression patterns were confirmed to an extent. An additional level of validation is provided through correlation with confirmed gene expression patterns published in the literature. *ANT* is present in the cotyledon up-regulated genes (Long and Barton, 1998), and an *AINTEGUMENTA*-like gene (At5g17430) present the root up-regulated genes, whose basal localisation was confirmed by promoter::GUS fusion (Casson *et al.*, 2005).

Functional analysis of the 50 most differentially expressed genes in the cotyledon and root provides a very different perspective to the temporal function analysis conducted for the apical and basal time-courses.

Basing predicted function for unknown genes on sequence similarity with the ~10% of the genome experimentally characterised, only ~30% of the genome remains without putative functional classification (The Arabidopsis Genome Initiative, 2000). Strikingly between 42% and 48% of the most up-regulated genes uncovered have no predicted function; however with a predicted spatial expression pattern during embryogenesis, analysis of these unknowns could reveal important functions.

The range of functions represented in the up-regulated root genes is limited compared to those of the cotyledon. The most significant functional group represented in the root is that of metabolism, this correlates well with the analysis performed by Yamada *et al.* (2003) on mature root tissue, which showed metabolism to be the largest functional group, comprising ~13% (plus 5.5% designated as protein metabolism) of the 549 root specific transcripts

identified. The cotyledon shows up-regulation of a wide range of functional groups, including energy and protein synthesis, this correlates with the increase in chloroplastic gene transcripts to reach a peak in the cotyledons of the mature embryo (Degenhardt *et al.*, 1991). A number of the genes functionally classified as protein synthesis are chloroplast and plastid ribosomal protein precursors. Knock-out mutants of plastid ribosomal proteins (S21 and L11) have been shown to impair photosynthesis and thus show an overlap with the up-regulation of energy function (Pesaresi *et al.*, 2001; Morita-Yamamuro *et al.*, 2004).

A significant number of putative transcription factors are present in both the cotyledon and root gene lists providing potential targets for further analysis into spatial control mechanisms.

An analysis of the overlap between significant apical and basal genes at each developmental stage was inconclusive as to the existence of a population of genes with a distinct spatial expression pattern throughout development. Expression pattern analysis of genes such as *PIN4* show that defined spatial expression along the embryonic time-course does exist (Friml, 2002). The statistical restrictions imposed by this analysis do not, however, reflect them as a statistically significant population. The analysis did show a significant overlap between the torpedo-stage of embryogenesis and the seedling, not just in root and cotyledon specific transcripts, but also in transcripts changing their spatial localisation. This may reflect a more consistent population of genes expressed in the late embryo whose maturation processes continue into the seedling, in contrast to early embryogenesis where more dynamic sets of genes may be required for pattern formation and morphogenesis. The existence of genes with changing spatial localisation through embryogenesis, and into the seedling, is highlighted by the small zinc finger-like protein (At5g50810/E) analysed in this work which changes from predominantly localised in the cotyledons of the torpedo-stage embryo to being predominantly localised in the root of the seedling.

Overall, the analysis of the data by GeneSpring reveals some potentially interesting findings; however it must be remembered that this analysis is a

starting point and not an end result. Many avenues of interest are revealed but further experimental research is required to both confirm and expand on these findings.

Chapter 6

Discussion

6.0 Discussion

6.1 Project objectives

As stated in Chapter 1, the overall objective of this thesis was to develop a protocol for the application of laser-capture microdissection to embryos of the model plant species *Arabidopsis thaliana*. Through the subsequent application of microarray technology and bioinformatic analysis, it was hoped to address the genetic control of apical-basal patterning.

This thesis describes the development of a protocol for using LCM to capture tissue from the cotyledon, root and shoot apical meristem regions of the torpedo stage embryo; utilisation of the *Arabidopsis* ATH1 GeneChip® to allow a global gene expression analysis and subsequent validation of data obtained using a variety of approaches; and further bioinformatic analysis to uncover potential genes of interest in spatial and temporal genetic control mechanisms.

6.2 Laser-mediated microdissection: A new tool for the high-resolution gene expression profiling of plant cells

The completion of genome sequencing programmes for many species has opened up new possibilities to investigate developmental and physiological events occurring throughout development, at a genomic level. Sequencing of both a representative dicot, *Arabidopsis thaliana* and a monocot, rice (*Oryza sativa*), has now been completed (Arabidopsis Genome Initiative, 2000; Goff *et al.*, 2003; Yu *et al.*, 2003). Commercially available oligonucleotide microarrays with significant genome coverage are available for both species making global analysis a realistic prospect for much of the plant research community.

However, as Girke *et al.* (2000) acknowledge current microarray technology does have its limitations particularly in the reliable detection of the most rarely expressed genes. High abundance transcripts have a significant dilution effect

on low abundance transcripts, and this effect is compounded by the use of heterogeneous tissue samples in which a population of transcripts may only be present in specific cell types. To bypass the dilution effect of heterogeneity sophisticated tools have been developed to isolate RNA from single cells or specific tissue types. One such tool is that of laser-mediated microdissection. Originally developed to allow the isolation of cancerous cells from complex tumour tissue (Emmert-Buck *et al.*, 1996), it has been used for many applications in the field of animal cell biology. Technical obstacles associated with plant cell architecture, has meant that this technology has only recently been adapted for use with plant tissue. The work described in this thesis forms part of this initial effort to develop protocols for application of this technology to plant tissue. In combination with other groups, protocols have now been developed for several plant species including *Arabidopsis*, maize and rice. RNA has been isolated from a range of tissue types and used for downstream applications including RT-PCR, cDNA library construction and microarray analysis. With the equipment required for laser-mediated microdissection widely available due to its use in pathology and cancer biology it is likely that its use in plant biology will continue to expand in the future.

6.3 Global gene expression profiling of plant tissue

In recent years the quantity of sequence data available for a range of organisms has increased rapidly. In combination with novel methodologies such as cDNA and oligonucleotide microarray, this wealth of data has permitted the simultaneous expression level analysis of thousands of genes in a single experiment, and given rise to the study of functional genomics (Brown and Botstein, 1999; Lockhart and Winzeler, 2000).

Examination of gene expression levels can reveal distinct temporal changes occurring over a developmental time-course, or molecular differences between different tissues or even individual cell types. Spatial expression profiling studies using both novel and traditional techniques for isolation of RNA from distinct organs or cell types have been successfully undertaken (Birnbaum *et al.*, 2003; Wellmer *et al.*, 2004; Lee *et al.*, 2005; Ma *et al.*, 2005).

Gene expression maps using microarray technology have been assembled for a number of species including *Caenorhabditis elegans* (Kim *et al.*, 2001) and *Drosophila melanogaster* (Stolc *et al.*, 2004). In *Arabidopsis thaliana* a number of large-scale projects have created data sets covering a range of organs and tissue types using such methods as full genome tiling arrays (Yamada *et al.*, 2003) and cell sorting of protoplasted root tissue in combination with the Affymetrix ATH1® array (Birnbaum *et al.*, 2003). The most comprehensive large-scale project AtGenExpress aimed to create an expression atlas for a wide range of tissue types and stages during the development of *Arabidopsis thaliana*. The first analysis of this data set has recently been published (Schmid *et al.*, 2005).

The work described in this thesis forms part of a rapidly growing collection of freely available data sets covering a diverse array of tissue types, developmental stages and physiological conditions. The general availability of this data provides a very powerful resource to the research community and should be of significant use for targeted experimental design.

A complementary technology for global gene expression analysis in defining the cellular specificity of expression patterns is that of large-scale *in situ* hybridisation. This has been utilised with great success in high throughput screens of *Drosophila* embryogenesis (Kopcynski *et al.*, 1998; Tomancak *et al.*, 2002) and the ascidian, *Ciona intestinalis* (Satou *et al.*, 2001; Imai *et al.*, 2004). Recently it has been applied to the development of the wheat (*Triticum aestivum*) caryopsis (Drea *et al.*, 2005). This study selected a set of 888 genes whose gene expression patterns were determined by microarray analysis to change during caryopsis development. Distinct novel spatial expression patterns were revealed, including those of a number of transcription factors spatially restricted to individual cell types. This study highlights the integration of high throughput technologies to provide candidates for further research.

The analysis of gene expression profiles does not however reveal the entire picture and must be treated with caution. Gygi *et al.* (1999) demonstrated that in yeast the mRNA abundance detected by global gene expression studies shows

a poor correlation with protein abundance, and therefore on its own, is not a sufficient measure by which to understand a biological system. For some genes exhibiting a constant gene expression level the protein abundance could differ by more than 20-fold, and conversely, constant protein abundance levels were observed for transcripts exhibiting differential gene expression of up to 30-fold (Gygi *et al.*, 1999). Such studies emphasise the importance of integrating transcriptional profiling with other global approaches such as proteomics and metabolomics in an integrated systems biology approach, in order to build a more complete picture of the biological system (Fiehn *et al.*, 2001; Provart and McCourt, 2004). One of the major challenges for systems biology in the future is how to integrate the wealth of data being produced in a coherent way so as to provide meaningful information to the general plant biologist. However, as the field of proteomics is not currently as advanced as that of transcriptomics, its application to studies such as the work described in this thesis are not currently technically feasible.

The validity of using sequence homology to mammals and yeast in order to predict the function and relative importance of plant genes, based on the assumption that fundamental gene expression mechanisms are entirely conserved between eukaryotes, is not entirely proven. Studies have shown that some genes predicted to be essential, based on other eukaryotes, display only mildly aberrant phenotypes when disrupted, and conversely, genes, which are dispensable in other organisms result in lethality when mutated in plants (Belostotsky and Rose, 2005). Therefore, it seems obvious that despite the powerful nature of these techniques, the conclusions reached should in many ways be treated as hypotheses, which require testing using traditional genetic and cell-biology methodologies, as hopefully has been demonstrated in this thesis (Provart and McCourt, 2004).

6.4 Final summary and future work

The objective of this thesis was to develop a protocol for the application of laser-capture microdissection to embryos of the model plant species *Arabidopsis thaliana*, which was undertaken with some success. Using this

protocol there are a number of further experimental directions of interest to explore.

As shown in Chapter 5 the use of only three developmental time-points was not conducive to the production of well defined clusters, ideally cotyledon and root tissue would be captured from the cotyledonary stage of embryogenesis to provide an additional time-point. The use of a user defined-model based clustering package could also have the potential to improve the groupings obtained.

Laser-capture microdissection could also be applied to mutant analysis. Preliminary work to capture the apical and basal regions of *hydra2/fackel* mutant embryos has been attempted without success using the current protocol, due to the histology being inadequate to enable reliable identification of the regions of interest. In order to take this work forwards the decision must be made to sacrifice RNA quality and quantity for the superior histology achievable using fixation and embedding protocols, such as the ethanol-acetic acid fixation and paraffin embedding proposed by Kerk *et al.* (2003).

The LCM system used in this work did not have the resolution to capture single cells and therefore areas of a defined size were captured. In the case of the torpedo-stage SAM this meant that there was the potential to capture a lot of non-specific material despite using the smallest beam size available. The use of alternative laser microdissection systems with the proposed ability to reliably capture smaller areas could allow the more accurate isolation of such tissue areas. The ability to capture smaller areas accurately would however potentially require an improvement in histology, only achievable through the use of detrimental procedures such as those described previously.

As well as the resource which is the LCM protocol for embryonic tissue, this work also created a huge amount of microarray data which has a great deal of potential for further use. The analysis of this data on its own or in combination with that of Casson *et al.* (2005), as described in this thesis, has uncovered a number of genes of potential interest including putative transcription factors showing spatial differential expression.

A number of putative transcription factors were selected from an initial analysis of the data for the purpose of validation and also further investigation. Promoter::GUS analysis of these genes revealed some interesting expression patterns, which validated the expression profiles observed in the microarray data and also added another level of specificity, such as for At5g14610 (construct C) which shows very specific localisation to the stomatal guard cells. An analysis of T-DNA insertion lines was conducted in an attempt to uncover a functional role for these genes, unfortunately no obvious aberrant phenotypes were observed. This could be due to a number of reasons, which could be addressed by additional work. The first is that the knockout phenotype could be conditional. Homozygous knockout plants could therefore be subjected to a range of physiologically challenging conditions. In addition to any aberrant phenotype observed the condition that induced it would also be highly instructive as to function. The second possibility is that the gene is functionally redundant as a result of membership of a gene-family, as transcription factor families are often relatively large there is the possibility of overlapping functions. To test for this a database search could be carried out to identify genes with high sequence homology to the one under investigation. Knockout lines should be obtained for these genes and double, triple and quadruple mutant crosses performed, any aberrant phenotype or additive aberrant phenotype resulting could then be related to the gene of interest. An example of such a study was described in Section 4.8.4, where Hua and Meyerowitz (1998) described functional redundancy in a family of putative ethylene receptors. Single mutants displayed no ethylene response defects, while a quadruple mutant manifested a constitutive ethylene triple response.

In addition to further work on the knockout phenotype, gain-of-function studies could also be conducted. By fusing the gene of interest to a strong constitutive promoter such as the CaMV35S promoter, the gene could be over-expressed in plants.

Any mutant phenotype of interest resulting from knockout or over-expression studies could then be analysed on a global scale by extracting RNA and conducting microarray experiments to deduce what transcripts are affected

compared to the wild-type and thus potential downstream targets for the transcription factor. Inducible systems have the potential to be very useful for time-courses and visualising directly regulated targets. The mutant (knockout) is initially transformed with the gene linked to an inducible promoter (alcohol, glucocorticoid or heat-shock). The gene is then induced along a time-course, linked to microarray analysis to determine which genes are directly affected.

In order to investigate the localisation of the protein encoded by the gene of interest, protein::GFP constructs could be assembled. Confocal microscopy could then be used to investigate the subcellular localisation of the protein of interest.

In addition to the putative transcription factors already under investigation, the GeneSpring analysis revealed a significant number of potentially interesting genes, for which T-DNA insertion lines could also be obtained to uncover putative functions.

This thesis accomplished its overall objective to develop a protocol for the application of laser-capture microdissection to embryos of the model plant species *Arabidopsis thaliana*. Additionally, through the subsequent application of microarray technology and bioinformatic analysis a resource has been created, which can be drawn upon to provide direction for future experimental work.

References

References

Aarts, M.G.M., Dirkse, W.G., Stiekema, W.J. and Pereira, A. (1993). Transposon tagging of a male sterility gene in *Arabidopsis*. *Nature* **363**, 715-717.

Affymetrix (2002). Statistical Algorithms Description Document http://www.affymetrix.com/support/technical/whitepapers/sadd_whitepaper.pdf

Aharoni, A. and Vorst, O. (2002). DNA microarrays for functional plant genomics. *Plant Molecular Biology* **48**, 99-118.

Aida, M., Ishida, T., Fukaka, H., Fujisawa, H. and Tasaka, M. (1997). Genes involved in organ separation in *Arabidopsis*: an analysis of the *cup-shaped cotyledon* mutant. *The Plant Cell* **9**, 841-857.

Aida, M., Ishida, T. and Tasaka, M. (1999). Shoot apical meristem and cotyledon formation during *Arabidopsis* embryogenesis: interaction among the *CUP-SHAPED COTYLEDON* and *SHOOT MERISTEMLESS* genes. *Development* **126**, 1563-1570.

Aida, M., Vernoux, T., Furutani, M., Traas, J. and Tasaka, M. (2002). Roles of *PIN-FORMED1* and *MONOPTEROS* in pattern formation of the apical region of the *Arabidopsis* embryo. *Development* **129**, 3965-3974.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. and Ecker, J.R. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653-657.

Apuya, N.R., Yadegari, R., Fischer, R.L., Harada, J.H. and Goldberg, R.B. (2002). *RASPBERRY3* gene encodes a novel protein important for embryo development. *Plant Physiology* **129**, 691-705.

Asano, T, Masumura, T., Kusano, H., Kikuchi, S., Shimada, H. and Kadowaki, K-I. (2002). Construction of a specialized cDNA library from plant cells isolated by laser capture microdissection: toward comprehensive analysis of the genes expressed in the rice phloem. *The Plant Journal* **32**, 401-408.

Aubourg, S., Kreis, M. and Lecharny, A. (1999). The DEAD box RNA helicase family in *Arabidopsis thaliana*. *Nucleic Acids Research* **27**, 628-636.

Ausubel, F. and Benfey, P. (2002). Arabidopsis functional genomics. *Plant Physiology* **129**, 393.

Ballare, C.L. (1999). Keeping up with the neighbours: phytochrome sensing and other signalling mechanisms. *Trends in Plant Science* **4**, 97-102.

Bancroft, I., Bhatt, A.M., Sjodin, C., Scofield, S., Jones, J.D.G. and Dean, C. (1992). Development of an efficient two-element transposon tagging system in *Arabidopsis thaliana*. *Molecular and General Genetics* **204**, 512-518.

Bancroft, I., Jones, J. D. G. and Dean, C. (1993). Heterologous transposon tagging of the *DRL1* locus in *Arabidopsis*. *The Plant Cell* **5**, 631-638.

Barton, K.M. and Poethig, R.S. (1993). Formation of the shoot apical meristem in *Arabidopsis thaliana*: an analysis of development in the wild type and in the shoot meristemless mutant. *Development* **119**, 823-831.

Baugh, L.R., Hill, A.A., Brown , E.L. and Hunter, C.P. (2001). Quantitative analysis of mRNA amplification by *in vitro* transcription. *Nucleic Acids Research* **29**, e29

Baumbusch, L.O., Sundal, I.K., Hughes, D.W., Galau, G.A. and Jakobsen, K.S. (2001). Efficient protocols for CAPS-based mapping in *Arabidopsis*. *Plant Molecular Biology Reporter* **19**, 137-149.

Bäumlein, H., Miséra, S., Luerßen, H., Kölle, K., Horstmann, C., Wobus, U. and Müller, A.J. (1994). The *FUS3* gene of *Arabidopsis thaliana* is a regulator of gene expression during late embryogenesis. *The Plant Journal* **6**, 379-387.

Bäurle, I. and Laux, T. (2003). Apical meristems: the plants fountain of youth. *BioEssays* **25**, 961-970.

Beemster, G.T.S., De Veylder, L., Vercruysse, S., West, G., Rombaut, D., Van Hummelen, P., Galichet, A., Gruissem, W., Inzé, D. and Vuylsteke, M. (2005). Genome-wide analysis of gene expression profiles associated with cell cycle transitions in growing organs of Arabidopsis. *Plant Physiology* **132**, 734-743.

Behringer, F.J. and Medford, J.I. (1992). A plasmid rescue technique for the recovery of plant DNA disrupted by T-DNA insertion. *Plant Molecular Biology Reporter* **10**, 190-198.

Belanger, K.D. and Quatrano, R.S. (2000a) Polarity: the role of localized secretion. *Current Opinion in Plant Biology* **3**, 67-72.

Belanger, K.D. and Quatrano, R.S. (2000b) Membrane recycling occurs during asymmetric tip growth and cell plate formation in *Fucus distichus* zygotes. *Protoplasma* **212**, 24-37.

Bellec, Y., Harrar, Y., Butaeye, C., Darnet, S., Bellini, C. and Faure, J.D. (2002). *Pasticcino2* is an protein tyrosine phosphatase-like involved in cell proliferation and differentiation in *Arabidopsis*. *The Plant Journal* **32**, 713-722.

Belostotsky, D.A. and Rose, A.B. (2005). Plant gene expression in the age of systems biology: integrating transcriptional and post-transcriptional events. *Trends in Plant Science* **10**, 347-535.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* **B57**, 289-300.

Benjamins, R., Quint, A., Weijers, D., Hooykaas, P. and Offringa, R. (2001). The PINOID protein kinase regulates organ development in *Arabidopsis* by enhancing polar auxin transport. *Development* **128**, 4057-4067.

Bennett, M.J., Marchant, A., Green, H.G., May, S.T., Ward, S.P., Millner, P.A., Walker, A.R., Schulz, B. and Feldmann, K.A. (1996). *Arabidopsis AUX1* gene: a permease-like regulator of root gravitropism. *Science* **273**, 948-950.

Berger, F., Taylor, A. and Brownlee, C. (1994). Cell fate determination by the cell wall in early *Fucus* development. *Science* **263**, 1421-1423.

Berleth, T. and Chatfield, S. (2002). Embryogenesis: Pattern formation from a single cell. In The Arabidopsis Book, C.R. Somerville and E.M. Meyerowitz, eds (Rockville: American Society of Plant Biologists). doi 10.1199/tab.0051, http://www.aspb.org/publications/arabidopsis/

Berleth, T. and Jürgens, G. (1993). The role of the *MONOPTEROS* gene in organising the basal body region of the *Arabidopsis* embryo. *Development* **118**, 575-587.

Bevan, M.W. (1984). Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Research* **12**, 8711-8721.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W. and Benfey, P.N. (2003). A gene expression map of the *Arabidopsis* root. *Science* **302**, 1956-1960.

Blilou, I., Frugier, F., Folmer, S., Serralbo, O., Willemsen, V., Wolkenfelt, H., Eloy, N.B., Ferreira, P.C.G., Weisbeek, P. and Scheres, B. (2002). The *Arabidopsis HOBBIT* gene encodes a CDC27 homolog that links the plant cell cycle to progression of cell differentiation. *Genes and Development* **16**, 2566-2575.

Blilou, I., Xu, J., Wildwater, M., Willemsen, V., Paponov, I., Friml, J., Heidstra, R., Aida, M., Palme, K. and Scheres, B. (2005). The PIN auxin efflux facilitator network controls growth and patterning in *Arabidopsis* roots. *Nature* **433**, 39-44.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M. and Benning, C. (1998). *AGO1* defines a novel locus of *Arabidopsis* controlling leaf development. *The EMBO Journal.* **17**, 170-180.

Bolstad, B.M. (2002). Comparing the effects of background, normalization and summarization on gene expression estimates. Unpublished manuscript (http://stat-www.berkeley.edu/users/bolstad/)

Bonner, R.F., Emmert-Buck, M.R., Cole, K., Pohida, T., Chuaqui, R.F., Goldstein, S.R. and Liotta, L.A. (1997). Cell sampling – laser capture microdissection: molecular analysis of tissue. *Science* **278**, 1481-1483.

Borroto, K.E. and Dure, L. III (1986). The expression of chloroplast genes during cotton embryogenesis. *Plant Molecular Biology* **7**, 105-113.

Boudet, N., Aubourg, S., Toffano-Nioche, C., Kreis, M. and Lecharny, A. (2001). Evolution of intron/exon structure of DEAD helicase family genes in *Arabidopsis*, *Caenorhabditis*, and *Drosophila*. *Genome Research* **11**, 2101-2114.

Bouget, F-Y., Gerttula, S., Shaw, S.L. and Quatrano, R.S. (1996). Localization of actin mRNA during the establishment of cell polarity and early cell divisions in *Fucus* embryos. *The Plant Cell* **8**, 189-201.

Boutilier, K., Offringa, R., Sharma, V.K., Kieft, H., Ouellet, T., Zhang, L., Hatton, J., Liu, C. –M., van Lammeren, A.A.M., Miki, B.L.A., Custers, J.B.M. and van Lookeren Campagne, M.M. (2002). Ectopic expression of *BABY BOOM* triggers a conversion from vegetative to embryonic growth. *The Plant Cell* **14**, 1737-1749.

Bowman, J.L., Alvarez, J., Weigel, D., Meyerowitz, E.M. and Smyth, D.R. (1993). Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. *Development* **119**, 721-743.

Brady, G. (2000). Expression profiling of single mammalian cells – small is beautiful. *Yeast* **17**, 211-217.

Brand, U., Fletcher, J.C., Hobe, M., Meyerowitz, E.M. and Simon, R. (2000). Dependence of stem cell fate in *Arabidopsis* on a feedback loop regulated by *CLV3* activity. *Science* **289**, 617-619.

Brandt, S., Kehr, J., Walz, C., Imlau, A., Willmitzer, L. and Fisahn, J. (1999). A rapid method for detection of plant gene transcripts from single epidermal, mesophyll and companion cells of intact leaves. *The Plant Journal* **20**, 245-250.

Brandt, S., Kloska, S., Altmann, T. and Kehr, J. (2002). Using array hybridisation to monitor gene expression at the single cell level. *Journal of Experimental Botany* **53**, 2315-2323.

Breton, C., Faure, J.E. and Dumas, S.C. (1995). From *in vitro* fertilization to early embryogenesis in maize. *Protoplasma* **187**, 3-12.

Brown, P.O. and Botstein, D. (1999). Exploring the new world of the genome with DNA microarrays. *Nature Genetics* **21**, 33-37.

Brownlee C. (2004). From Polarity to pattern: early development in fucoid algae. In Polarity of Plants, K. Lindsey, ed(Oxford: Blackwell Publishing), pp 138-156.

Brownlee, C. and Berger, F. (1995). Developmental signals and their transduction during *Fucus* embryogenesis. *Progress in Phycological Research* **11**, 103-144.

Brownlee, C. and Bouget, F-Y. (1998). Polarity determination in *Fucus*: from zygote to multicellular embryo. *Seminars in Cell and Developmental Biology* **9**, 179-185.

Brunner, A.M., Yakovlev, I.A. and Strauss, S.H. (2004). Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biology* **4**, 14.

Burgess, J.K. (2001). Gene expression studies using microarrays. *Clinical and Experimental Pharmacology and Physiology* **28**, 321-328.

Busch, M., Mayer, U. and Jürgens, G. (1996). Molecular analysis of the *Arabidopsis* pattern formation gene *GNOM*: gene structure and intragenic complementation. *Molecular and General Genetics* **250**, 681-691.

Byrne, M.E., Barley, R., Curtis, M., Arroyo, J.M., Dunham, M., Hudson, A. and Martienssen, R.A. (2000). *Asymmetric leaves1* mediates leaf patterning and stem cell function in *Arabidopsis*. *Nature* **408**, 967-971.

Casson, S., Spencer, M., Walker, K. and Lindsey, K. (2005). Laser capture microdissection for the analysis of gene expression during embryogenesis of *Arabidopsis*. *The Plant Journal* **42**, 111-123.

Chang, Y.C. and Walling, L.L. (1991). Abscisic acid negatively regulates expression of chlorophyll *a/b* binding protein genes during soybean development. *Plant Physiology* **97**, 1260-1264.

Chang, Y.C. and Walling, L.L. (1992). Chlorophyll *a/b* binding protein genes are differentially expressed during soybean development. *Plant Molecular Biology* **19**, 217-230.

Chilton, M.-D., Drummond, M.H., Merlo, D.J., Sciaky, D., Montoya, A.L., Gordon, M.P. and Nester, E.W. (1977). Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. *Cell* **11**, 263-271.

Christensen, S.K., Dagenais, N., Chory, J. and Weigel, D. (2000). Regulation of auxin response by the protein kinase PINOID. *Cell* **100**, 469-478.

Christianson, M.L. (1986). Fate map of the organizing shoot apex in *Gossypium*. *American Journal of Botany* **73**, 947-958.

Cho, R.J., Huang, M., Campbell, M.J., Dong, H., Steinmetz, L., Sapinoso, G.H., Elledge, S.J., Davis, R.W. and Lockhart, D.J. (2001). Transcriptional regulation and function during the human cell cycle. *Nature Genetics* **27**, 48-54.

Clare, A. and King, R.D. (2002). How well do we understand the clusters found in microarray data? *In Silico Biology* **2**, 511-522.

Clark, S.E. (2001). Meristems: start your signalling. *Current Opinion in Plant Biology* **4**, 28-32.

Clark, S.E., Running, M.P. and Meyerowitz, E.M. (1993). *CLAVATA1*, a regulator of meristem and flower development in *Arabidopsis*. *Development* **119**, 397-418.

Clark, S.E., Running, M.P. and Meyerowitz, E.M. (1995). *CLAVATA3* is a specific regulator of shoot and floral meristem development affecting the same processes as *CLAVATA1*. *Development* **121**, 2057-2067.

Clark, S.E., Jacobsen, S.E., Levin, J.Z. and Meyerowitz, E.M. (1996). The *CLAVATA* and *SHOOT MERISTEMLESS* loci competitively regulate meristem activity in *Arabidopsis*. *Development* **122**, 1567-1575.

Clough, S.J. and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium* - mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**, 735-743.

Coen, E.S., Robbins, T.P., Almeida, J., Hudson, A. and Carpenter, R. (1989). Consequences and mechanisms of transposition in *Antirrhinum majus*. In *Mobile DNA*, D.E. Berg, M.M. Howe, eds(Washington DC: *American Society for Microbiology*). pp. 413-436.

Conway, L.J. and Poethig, R.S. (1997). Mutations of *Arabidopsis thaliana* that transform leaves into cotyledons. *Proceedings of the National Academy of Sciences USA* **94**, 10209-10214.

Cope, L.M., Irizarry, R.A., Jaffee, H.A., Wu, Z. and Speed, T.P. (2004). A benchmark for Affymetrix GeneChip expression measures. *Bioinformatics* **20**, 323-331.

Cosgrove, D.J. (2000). Loosening of plant cell walls by expansins. *Nature* **407**, 321-326.

Craigon, D.J., James, N., Okyere, J., Higgins, J., Jotham, J. and May, S. (2004). NASCArrays: a repository for microarray data generated by NASC's Transcriptomics Service. *Nucleic Acids Research* **32** (Database issue), D575-D577.

Davuluri, R.V., Sun, H., Palaniswamy, S.K., Matthews, N., Molina, C., Kurtz, M. and Grotewold, E. (2003). AGRIS: Arabidopsis Gene Regulatory Information Server, an information resource of Arabidopsis *cis*-regulatory elements and transcription factors. *BMC Bioinformatics* **4**, 25.

Dean, C., Sjodin, C., Page, T., Jones, J. and Lister, C. (1992). Behaviour of the maize transposable element *Ac* in *Arabidopsis thaliana*. *The Plant Journal* **2**, 69-81.

Degenhardt, J., Fiebig, C. and Link, G. (1991). Chloroplast and nuclear transcripts for plastid proteins in *Arabidopsis thaliana*: tissue distribution in mature plants and during seedling development and embryogenesis. *Botanica Acta* **104**, 455-463.

DeRisi, J.L., Iyer, V.R. and Brown, P.O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680-686.

Dharmasiri, N., Dharmasiri, S. and Estelle, M. (2005). The F-box protein TIR1 is an auxin receptor. *Nature* 441-445.

Doelling, J.H., Yan, N., Kurepa, J., Walker, J. and Vierstra, R.D. (2001). The ubiquitin-specific protease UBP14 is essential for early embryo developmental in *Arabidopsis thaliana*. *The Plant Journal* **27**, 393-405.

Dolan, L., Janmaat, K., Willemsen, V., Linstead, P., Poethig, S., Roberts, K., and Scheres, B. (1993). Cellular Organization of the *Arabidopsis thaliana* root. *Development* **119**, 71-84.

Drea, S., Leader, D.J., Arnold, B.C., Shaw, P., Dolan, L. and Doonan, J.H. (2005). Systematic spatial analysis of gene expression during wheat caryopsis development. *The Plant Cell* **17**, 2172-2185.

Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P. and Trent, J.M. (1999). Expression profiling using cDNA microarrays. *Nature Genetics* **21**, 10-14.

Dure, L. III (1993). A repeating 11-mer amino acid motif and plant desiccation. *The Plant Journal* **3**, 363-369.

Edwards, K., Johnstone, C. and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research* **19**, 1349.

Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences USA* **95**, 14863-14868.

Elliot, R.C., Betzner, A.S., Huttner, E., Oakes, M.P., Tucker, W.Q., Gerentes, D., Perez, P. and Smyth, D.R. (1996). *AINTEGUMENTA*, an *APETELA-2* like gene in *Arabidopsis* with pleiotropic roles in ovule development and floral organ growth. *The Plant Cell* **8**, 155-168.

Emmert-Buck, M.R., Bonner, R.F., Smith, P.D., Chuaqui, R.F., Zhuang, Z.P., Goldstein, S.R., Weiss, R.A. and Liotta, L.A. (1996). Laser capture microdissection. *Science* **274**, 998-1001.

Eshed, Y., Baum, S.F., Perea, J.V. and Bowman, J.L. (2001). Establishment of polarity in lateral organs of plants. *Current Biology* **11**, 1251-1260.

Faure, J.-D., Vittorioso, P., Santoni, V., Fraisier, V., Prinsen, E., Barlier, I., Van Onckelen, H., Caboche, M. and Bellini, C. (1998). The *PASTICCINO* genes of *Arabidopsis thaliana* are involved in the control of cell division and differentiation. *Development* **125**, 909-918.

Fiehn, O., Kloska, S. and Altmann, T. (2001). Integrated studies on plant biology using multiparallel techniques. *Current Opinion in Biotechnology* **12**, 82-86.

Fisher, R.H., Barton, M.K., Cohen, J.D. and Cooke, T.J. (1996). Hormonal studies of *fass*, an *Arabidopsis* mutant that is altered in organ elongation. *Plant Physiology* **110**, 1109-1121.

Fleming, A.J. (2005). Formation of primordial and phyllotaxy. *Current Opinion in Plant Biology* **8**, 53-58.

Fletcher, J.C. (2002). Shoot and floral meristem maintenance in Arabidopsis. *Annual Review of Plant Biology* **53**, 45-66.

Fletcher, J.C., Brand, U., Running, M.P., Simon, R. and Meyerowitz, E.M. (1999). Signalling of cell fate decision by CLAVATA-3 in *Arabidopsis* shoot meristems. *Science* **283**, 1911-1914.

Foder, S.P., Read, J.L., Pirrung, M.C., Stryer, L., Lu, A.T. and Solas, D. (1991). Light-directed, spatially addressable parallel chemical synthesis. *Science* **251**, 767-773.

Franzmann, L.H., Yoon, E.S. and Meinke, D.W. (1995). Saturating the genetic map of *Arabidopsis thaliana* with embryonic mutations. *The Plant Journal* **7**, 341-350.

Friml, J., Benková, E., Blilou, I., Wisniewska, J., Hamann, T., Ljung, K., Woody, S., Sandberg, G., Scheres, B., Jürgens, G and Palme, K. (2002). AtPIN4 mediates sink-driven auxin gradients and root patterning in *Arabidopsis*. *Cell* **108**, 661-673.

Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwartz, H., Hamann, T., Offriga, R. and Jürgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. *Nature* **426**, 147-153.

Friml, J., Yang, X., Michniewicz, M., weijers, D., Quint, A., Tietz, O., Benjamins, R., Ouwerkerk, P.B.F., Ljung, K., Sandberg, G., Hooykaas, P.J.J., Palme, K. and Offringa, R. (2004). A PINOID-dependent binary switch in apical-basal PIN polar targeting directs auxin efflux. *Science* **306**, 862-865.

Galbraith, D.W., Harkins, K.R. and Jefferson, R.A. (1988). Flow cytometric characterization of the chlorophyll contents and size distributions of plant protoplasts. *Cytometry* **9**, 75-83.

Galweiler, L., Guan, C., Muller, A., Wisman, E., Mendgen, K., Yephremov, A. and Palme, K. (1998). Regulation of polar auxin transport by *AtPIN1* in *Arabidopsis* vascular tissue. *Science* **282**, 2226-2230.

Gasser, C.S., Broadhvest, B., and Hauser, B.A. (1998). Genetic analysis of ovule development. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**, 1-24.

Gaudin, V., Vrain, T. and Jouanin, L. (1994). Bacterial genes modifying hormonal balances in plants. *Plant Physiology and Biochemistry* **32**, 11-29.

Geldner, N. (2004). The plant endosomal system – its structure and role in signal transduction and plant development. *Planta* **219**, 547-560.

Geldner, N., Anders, N., Wolters, H., Keicher, J., Kornberger, W., Muller, P., Delbarre, A., Ueda, T., Nakano, A. and Jürgens, G. (2003). The *Arabidopsis* GNOM ARF-GEF mediates endosomal recycling, auxin transport, and auxin-dependent plant growth. *Cell* **112**, 219-230.

Geldner, N., Friml, J., Stierhof, Y.-D., Jürgens, G. and Palme, K. (2001). Auxin transport inhibitors block PIN1 cycling and vesicle trafficking. *Nature* **413**, 425-428.

Gerats, A.G.M., Huits, H., Vrijlandt, E., Maraña, C., Souer, E. and Beld, M. (1990). Molecular characterization of a non-autonomous transposable element (*dTph1*) of petunia. *The Plant Cell* **2**, 1121-1128.

Gheysen, G., Angenon, G. and Van Montagu, M. (1998). *Agrobacterium*-mediated plant transformation: a scientifically intriguing story with significant applications. *Transgenic Plant Research* **7**, 1-33.

Gheysen, G., Villaroel, R. and Van Montagu, M. (1991). Illegitimate recombination in plants: a model for T-DNA integration. *Genes and Development* **5**, 287-297.

Gillespie, J.W., Ahram, M., Best, C.J., Swalwell, J.I., Krizman, D.B., Petricoin, E.F., Liotta, L.A. and Emmert-Buck, M.R. (2001). The role of tissue microdissection in cancer research. *Cancer Journal* **7**, 32-39.

Gillespie, J.W., Best, C.J.M., Bischel, V.E., Cole, K.A., Greenhut, S.F., Hewitt, S.M., Ahram, M., Gathright, Y.B., Merino, M.J., Strausberg, R.L., Epstein, J.I., Hamilton, S.R., Gannot, G., Baibakova, G.V., Calvert, V.S., Flaig, M.J., Chuaqui, R.F., Herring, J.C., Pfeifer, J., Petricoin, E.F., Linehan, W.M., Duray, P.H., Bova, G.S. and Emmert-Buck, M.R. (2002). Evaluation of non-formalin tissue fixation for molecular profiling studies. *American Journal of Pathology* **160**, 449-457.

Girke, T., Todd, J., Ruuska, S., Benning, C. and Ohlrogge, J. (2000). Microarray analysis of developing Arabidopsis seeds. *Plant Physiology* **124**, 1570-1581.

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H. Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.-L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A. and Briggs, S. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92-100.

Goldberg, R.B., Barker, S.J. and Perez-Grau, L. (1989). Regulation of gene expression during plant embryogenesis. *Cell* **56**, 149-160.

Goldberg, R.B. and Harada, J.J. (1994). *LEAFY COTYLEDON1* is an essential regulator of late embryogenesis and cotyledon identity in *Arabidopsis*. *The Plant Cell* **6**, 1731-1745.

Goldsworthy, S.M., Stockton, P.S., Trempus, C.S., Foley, J.F. and Maronpot, R.R. (1999). Effects of fixation on RNA extraction and amplification from laser capture microdissected tissue. *Molecular Carcinogenesis* **25**, 86-91.

Goode, B.L., Drubin, D.G. and Barnes, G. (2000). Functional cooperation between the microtubule and actin cytoskeletons. *Current Opinion in Cell Biology* **12**, 63-71.

Grant, S.G.N., Jessee, J., Bloom, F.R. and Hanahan, D. (1990). Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proceedings of the National Academy of Sciences USA*. **87**, 4645-4649.

Gray, W.M., Kepinski, S., Rouse, D. Leyser, O., and Estelle, M. (2001). Auxin regulates SCF^{TIR1}-dependent degradation of AUX/IAA proteins. *Nature* **414**, 271–276.

Grosset, J., Marty, I., Chartier, Y. and Meyer, Y. (1990). mRNA's newly synthesized by tobacco mesophyll protoplasts are wound-inducible. *Plant Molecular Biology* **15**, 485-496.

Gruenberg, J. (2001). The endocytic pathway: a mosaic of domains. *Nature Reviews Molecular Cell Biology* **2**, 721-730.

Guilfoyle, T., Hagen, G., Ulmasov, T. and Murfett, J. (1998). How does Auxin turn on genes? *Plant Physiology* **118**, 341-347.

Gygi, S.P., Rochon, Y., Franza, B.R. and Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. *Molecular and Cellular Biology* **19**, 1720-1730.

Haberer, G., Erschadi, S. and Torres Ruiz, R.A. (2002). The *Arabidopsis* gene *PEPINO/PASTICCINO2* is required for proliferation control of meristematic and non-meristematic cells and encodes a putative anti-phosphatase. *Development Genes and Evolution* **212**, 542-550.

Hable, W.E. and Kropf, D.L. (2000) Sperm entry induces polarity in fucoid zygotes. *Development* **127**, 493-501.

Hadfi, K., Speth, V. and Neuhaus, G. (1998). Auxin-induced developmental patterns in *Brassica juncea* embryos. *Development* **125**, 879-887.

Haecker, A., Groß-Hardt, R., Geiges, B., Sarkar, A., Breuninger, H., Herrmann, M. and Laux, T. (2004). Expression dynamics of *WOX* genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*. *Development* **131**, 657-668.

Hamann, T., Benkova, E., Bäurle, I., Kientz, M. and Jürgens, G. (2002). The *Arabidopsis BODENLOS* gene encodes an auxin response protein inhibiting *MONOPTEROS*-mediated embryo patterning. *Genes and Development* **16**, 1610-1615.

Hamann, T., Mayer, U. and Jürgens, G. (1999). The auxin-insensitive *bodenlos* mutation affects primary root formation and apical-basal patterning in the *Arabidopsis* embryo. *Development* **126**, 1387-1395.

Hansel, D.E., Rahman, A., Hidalgo, M., Thuluvath, P.J., Lillemoe, K.D., Schulick, R., Ku, J.L., Park, J.G., Miyazaki, K., Ashfaq, R., Wistuba, I.I., Varma, R., Hawwthorne, L., Geradts, J., Argani, P. and Maitra, A. (2003). Identification of novel cellular targets in biliary tract cancers using global gene expression technology. *American Journal of Pathology* **163**, 217-229.

Hardtke, C.S. and Berleth, T. (1998). The *Arabidopsis* gene *MONOPTEROS* encodes a transcription factor mediating embryo axis formation and vascular development. *The EMBO Journal* **17**, 1405-1411.

Harrington, C.A., Rosenow, C. and Retief, J. (2000). Monitoring gene expression using DNA microarrays. *Current Opinion in Microbiology* **3**, 285-291.

Hawes, M.C. and Smith, L.Y. (1989). Requirement for chemotaxis in pathogenicity of *Agrobacterium tumefaciens* on roots of soil-grown pea plants. *Journal of Bacteriology*, **171**, 5668-5671.

He, Y.D., Dai, H., Schadt, E.E., Cavet, G., Edwards, S.W., Stepaniants, S.B., Duenwald, S., Kleinhanz, R., Jones, A.R., Shoemaker, D.D. and Stoughton, R.B. (2003). Microarray standard data set and figures of merit for comparing data processing methods and experiment designs. *Bioinformatics* **19**, 956-965.

Hellmann, H., Hobbie, L., Chapman, A., Dharmasiri, S., Dharmasiri, N., del Pozo, C., Reinhardt, D. and Estelle, M. (2003). *Arabidopsis AXR6* encodes CUL1 implicating SCF E3 ligases in auxin regulation of embryogenesis. *The EMBO Journal* **22**, 3314-3325.

Henikoff, S., Greene, E.A., Pietrokovski, S., Bork, P., Attwood, T.E. and Hood, L. (1997). Genome families: the taxonomy of protein paralogs and chimeras. *Science* **278**, 609-614.

Hennig, L., Gruissem, W., Grossniklaus, U. and Köhler, C. (2004). Transcriptional programs of early reproductive stages in Arabidopsis. *Plant Physiology* **135**, 1765-1775.

Hilson, P., Small, I. and Kuiper, M.T.R. (2003). European consortia building integrated resources for *Arabidopsis* functional genomics. *Current Opinion in Plant Biology* **6**, 426-429.

Hirner, B., Fischer, W.N., Rentsch, D., Kwart, M. and Frommer, W.B. (1998). Developmental control of H+/amino acid permease gene expression during seed development of *Arabidopsis*. *The Plant Journal* **14**, 535-544.

Hirsch, R.E., Lewis, B.D., Spalding, E.P. and Sussman, M.R. (1998). A role for the AKT1 potassium channel in plant nutrition. *Science* **280**, 918-921.

Hobbie, L., McGovern, M., Hurwitz, L.R., Pierro, A., Lui, N.Y., Bandyopadhyay, A. and Estelle, M. (2000). The *axr6* mutants of *Arabidopsis thaliana* define a gene involved in auxin response and early development. *Development* **127**, 23-32.

Hoekema, A., Hirsch, P.R., Hooykaas, P.J.J. and Schilperoort, R.A. (1983). A binary plant vector strategy based on separation of *vir-* and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature* **303**, 179-180.

Hong, S.K., Kitano, H., Satoh, H. and Nagoto, Y. (1996). How is embryo size genetically regulated in rice? *Development* **122**, 2051-2058.

Hua, J. and Meyerowitz, E.M. (1998). Ethylene responses are negatively regulated by a receptor gene family in *Arabidopsis thaliana*. *Cell* **94**, 261-271.

Hughes, T.R., Marton, M.J., Jones, A.R., Roberts, C.J., Stoughton, R., Armour, C.D., Bennett, H.A., Coffey, E., Dai, H., He, Y.D., Kidd, M.J., King, A.M., Meyer, M.R., Slade, D., Lum, P.Y., Stepaniants, S.B., Shoemaker, D.D., Gachotte, D., Chakraburtty, Simon, J., Bard, M. and Friend, S.H. (2000). Functional discovery via a compendium of expression profiles. *Cell* **102**, 109-126.

Imai, K.S., Hino, K., Yagi, K., Satoh, N. and Satou, Y. (2004). Gene expression profiles of transcription factors and signalling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. *Development* **131**, 4047-4058.

Inada, N. and Wildermuth, M.C. (2005). Novel tissue preparation method and cell-specific marker for laser microdissection of *Arabidopsis* mature leaf. *Planta* **221**, 9-16.

Ingham, P.W. (1988). The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature* **335**, 25-34.

Irish, V. and Sussex, I.M. (1992). A fate map of the *Arabidopsis* embryonic shoot apical meristem. *Development* **115**, 745-753.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B. and Speed, T.P. (2003a). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research* **31**, e15.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U. and Speed, T.P. (2003b). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264.

Iscove, N.N., Barbara, M., Gu, M., Gibson, M., Modi, C. and Winegarden, N. (2002). Representation is faithfully preserved in global cDNA amplified from sub-picogram quantities of mRNA. *Nature Biotechnology* **20**, 940-943.

Ishida, T., Aida, M., Takada, S. and Tasaka, M. (2000). Involvement of CUP-SHAPED COTYLEDON genes in gynoecium and ovule development in *Arabidopsis thaliana. Plant Cell Physiology.* **41**, 60-67.

Iwakawa, H., Ueno, Y., Semiarti, E., Onouchi, H., Kojima, S., Tsukaya, H., Hasebe, M., Soma, T., Ikezaki, M., Machida, C. and Machida, Y. (2002). The *ASYMMETRIC LEAVES2* gene of *Arabidopsis thaliana*, required for formation of a symmetric flat leaf lamina, encodes a member of a novel family of proteins characterized by cysteine repeats and a leucine zipper. *Plant Cell Physiology* **5**, 467-478.

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3**, 318-356.

Jang, J.-C., Fulioka, S., Tasaka, M., Seto, H., Takatsuto, S., Ishii, A., Aida, M., Yoshida, S. and Sheen, J. (2000). A critical role of sterols in embryonic patterning and meristem programming revealed by the *fackel* mutants of *Arabidopsis thaliana. Genes and Development* **14**, 1485-1497.

Jiang, K. and Feldman, L.J. (2005). Regulation of root apical meristem development. *Annual Review of Cell and Developmental Biology* **21**, 485-509.

Jaffe, L.F. (1958). Tropistic responses of zygotes of the fucaceae to polarised light. *Experimental Cell Research* **15**, 282-299.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987). GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO Journal* **6**, 3901-3907.

Jenik, P.D. and Barton, M.K. (2005). Surge and destroy: the role of auxin in plant embryogenesis. *Development* **132**, 3577-3585.

Jones, L. and McQueen-Mason, S. (2004). A role for expansins in dehydration and rehydration of the resurrection plant *Craterostigma plantagineum. FEBS Letters* **559**, 61-65.

Jürgens, G. (2001). Apical-basal pattern formation in *Arabidopsis* embryogenesis. *The EMBO Journal* **20**, 3609-3616.

Jürgens, G., Torres Ruiz, R.A. and Berleth, T. (1994). Embryonic pattern formation in flowering plants. *Annual Review of Genetics* **28**, 351-371.

Kane, M.D., Jatkoe, T.A., Stumpf, C.R., Lu, J., Thomas, J.D. and Madore, S.J. (2000). Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Research* **28**, 4552-4557.

Kaplan, D.R. and Cooke, T.J. (1997). Fundamental concepts in the embryogenesis of dicotyledons: a morphological interpretation of embryo mutants. *The Plant Cell* **9**, 1903-1919.

Karrer, E.E., Lincoln, J.E., Hogenhout, S., Bennett, A.B., Bostock, R.M., Martineau, B., Lucas, W.J., Gilchrist, D.G. and Alexander, D. (1995). *In situ* isolation of mRNA from individual plant cells: creation of cell-specific cDNA libraries. *Proceedings of the National Academy of Sciences USA* **92**, 3814-3818.

Keith, K., Kraml, M., Dengler, N.G. and McCourt, P. (1994). *FUSCA3*: A heterochronic mutation affecting late embryo development in *Arabidopsis*. *The Plant Cell* **6**, 589-600.

Kempin, S.A., Savidge, B. and Yanofsky, M.F. (1995). Molecular basis of the cauliflower phenotype in *Arabidopsis*. *Science* **267**, 522-525.

Kepinski, S. and Leyser, O. (2005). The *Arabidopsis* F-box protein TIR1 is an auxin receptor. *Nature* **435**, 446-451.

Kerk, N.M., Ceserani, T., Tausta, S.L., Sussex, I.M. and Nelson, T.M. (2003). Laser capture microdissection of cells from plant tissues. *Plant Physiology* **132**, 27-35.

Kermode, A.R. (1990). Regulatory mechanisms involved in the transition from seed development to germination. *CRC Critical Reviews in Plant Science* **81**, 280-288.

Kerstetter, R.A., Bollman, K., Tayer, R.A., Bomblies, K. and Poethig, S. (2001). *KANADI* regulates organ polarity in *Arabidopsis*. *Nature* **411**, 706-709.

Kim, J., Harter, K., and Theologis, A. (1997). Protein–protein interactions among the Aux/IAA proteins. *Proceedings of the National Academy of Sciences USA* **94**, 11786–11791.

Kim, S.K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J.M., Eizinger, A., Wylie, B.N. and Davidson, G.S. (2001). A gene expression map for *Caenorhabditis elegans*. *Science* **293**, 2087-2092.

King, C., Guo, N., Frampton, G.M., Gerry, N.P., Lenburg, M.E. and Rosenberg, C.L. (2005). Reliability and reproducibility of gene expression measurements using amplified RNA from laser-microdissected primary breast tissue with oligonucleotide arrays. *The Journal of Molecular Diagnostics* **7**, 57-64.

King, R.C. and Stansfield, W.D. (1997). A Dictionary of Genetics. 5th ed (New York: Oxford University Press).

Kirk, M.M., Ransick, A., McRae, S.E. and Kirk, D.L. (1993). The relationship between cell size and cell fate in *Volvox carteri*. *The Journal of Cell Biology* **123**, 191-208.

Klucher, K.M., Chow, H., Reiser, L. and Fisher, R.L. (1996). The *AINTEGUMENTA* gene of Arabidopsis required for ovule and female gametophyte development is related to the floral homeotic gene *APETALA2*. *The Plant Cell* **8**, 137-153.

Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z., Körber, H., Redei, G.P. and Schell, J. (1989). High-frequency T-DNA mediated gene tagging in plants. *Proceedings of the National Academy of Sciences USA* **86**, 8467-8471.

Koncz, C. and Schell, J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. *Molecular and General Genetics* **204**, 383-396.

Kopczak, S.D., Haas, N.A., Hussey, P.J., Silflow, C.D. and Snustad, D.P. (1992). The small genome of Arabidopsis contains at least six expressed α-tubulin genes. *The Plant Cell* **4**, 539-547.

Kopczynski, C.C., Noordermeer, J.N., Serano, T.L., Chen, W.-Y., Pendleton, J.D., Lewis, S., Goodman, C.S. and Rubin, G.M. (1998). A high throughput screen to identify secreted and transmembrane involved in *Drosophila* embryogenesis. *Proceedings of the National Academy of Sciences USA* **95**, 9973-9978.

Kropf, D.L. (1997). Induction of polarity in fucoid zygotes. *The Plant Cell* **9**, 1011-1020.

Kropf, D.L., Bisgrove, S.R. and Hable, W.E. (1999). Establishing a growth axis in fucoid algae. *Trends in Plant Science* **4**, 490-494.

Krysan, P.J., Young, J.C. and Sussman, M.R. (1999). T-DNA as an insertional mutagen in Arabidopsis. *The Plant Cell* **11**, 2283-2290.

Krysan, P.J., Young, J.C., Tax, F. and Sussman, M.R. (1996). Identification of transferred DNA insertions within *Arabidopsis* genes involved in signal transduction and ion transport. *Proceedings of the National Academy of Sciences. USA* **93**, 8145-8150.

Kumaran, M.K., Bowman, J.L. and Sundaresan, V. (2002). *YABBY* polarity genes mediate the repression of *KNOX* homeobox genes in *Arabidopsis*. *The Plant Cell* **14**, 2761-2770.

Laux, T. and Jürgens, G. (1997). Embryogenesis: a new start in life. *The Plant Cell* **9**, 989-1000.

Laux, T. and Mayer, K.F.X. (1998). Cell fate regulation in the shoot meristem. *Seminars in Cell and Developmental Biology* **9**, 195-200.

Laux, T., Mayer, K.F.X., Berger, J. and Jürgens, G. (1996). The *WUSCHEL* gene is required for shoot and floral meristem integrity in *Arabidopsis*. *Development* **122**, 87-96.

Lee, J.-Y., Levesque, M. and Benfey, P.N. (2005). High-throughput RNA isolation technologies. New tools for high-resolution gene expression profiling in plant systems. *Plant Physiology* **138**, 585-590.

Lenhard, M. and Laux, T. (1999). Shoot meristem formation and maintenance. *Current Opinion in Plant Biology* **2**, 44-50.

Lennon, G.G. and Lehrach, H. (1991). Hybridization analyses of arrayed cDNA libraries. *Trends in Genetics* **7**, 314-317.

Leyser, O. and Day, S. (2003). Mechanisms in Plant development. (Oxford: Blackwell Publishing).

Li, Y., Jones, L. and McQueen-Mason, S. (2003). Expansins and cell growth. *Current Opinion in Plant Biology* **6**, 603-610.

Lindsey, K. and Topping, J.F. (1993). Embryogenesis: a question of pattern. *Journal of Experimental Botany* **44**, 359-374.

Lipshutz, R.J., Fodor, S.P., Gingeras, T.R. and Lockhart, D.J. (1999). High density synthetic oligonucleotide arrays. *Nature genetics* **21**, 20-24.

Liu, C.M. McElver, J., Tzafrir, I., Joosen, R., Wittich, P., Patton, D., Van Lammeren, A.A.M. and Meinke, D. (2002). Condensin and cohesin knockouts in *Arabidopsis* exhibit a *titan* seed phenotype. *The Plant Journal* **29**, 405-415.

Liu, C.M. and Meinke, D.W. (1998). The *titan* mutants of *Arabidopsis* are disrupted in mitosis and cell cycle control during seed development. *The Plant Journal* **16**, 21-31.

Liu, C.M., Xu, Z.H. and Chua, N.H. (1993). Auxin polar transport is essential for the establishment of bilateral symmetry during early plant embryogenesis. *The Plant Cell* **5**, 621-630.

Liu, X., Wang, H., Li, Y., Tang, Y., Liu, Y., Hu, X., Jia, P., Ying, K., Feng, Q., Guan, J., Jin, C., Zhang, L., Lou, L., Zhou, Z. and Han, B. (2004). Preparation of single rice chromosome for construction of a DNA library using a laser microbeam trap. *Journal of Biotechnology* **109**, 217-226.

Lockhart, D.J. and Winzeler, E.A. (2000). Genomics, gene expression and DNA arrays. *Nature* **405**, 827-836.

Long, J.A. and Barton, M.K. (1998). The development of apical embryonic pattern in *Arabidopsis. Development* **125**, 3027-3035.

Long, J. and Barton, M.K. (2000). Initiation of axillary and floral meristems in *Arabidopsis. Developmental Biology* **218**, 341-353.

Long, J.A., Moan, E.I., Medford, J.I. and Barton, M.K. (1996). A member of the *KNOTTED* class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis. Nature* **379**, 66-69.

Lopes, M.A. and Larkins, B.A. (1993). Endosperm origin, development and function. *The Plant Cell* **5**, 1383-1399.

Lotan, T., Ohto, M., Yee, K.M., West, M.A.L., Lo, R., Kwong, R.W., Yamagishi, K., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (1998). *Arabidopsis LEAFY COTYLEDON1* is sufficient to induce embryo development in vegetative cells. *Cell* **93**, 1195-1205.

Lu, P., Porat, R., Nadeau, J.A. and O'Neill, S.D. (1996). Identification of a meristem L1 layer-specific gene in *Arabidopsis* that is expressed during embryonic pattern formation and defines a new class of homeobox genes. *The Plant Cell* **8**, 2155-2168.

Lukowitz, W., Roeder, A., Parmenter, D. and Somerville, C. (2004). A MAPKK kinase gene regulates extra-embryonic cell fate in *Arabidopsis. Cell* **116**, 109-119.

Luo, Y. and Koop, H.U. (1997). Somatic embryogenesis in cultured immature zygotic embryos and leaf protoplasts of *Arabidopsis thaliana. Planta* **202**, 387-396.

Lyczak, R., Gomes, J.E. and Bowerman, B. (2002). Heads or tails: cell polarity and axis formation in the early *Caenorhabditis elegans* embryo. *Developmental Cell* **3**, 157-166.

Lynn, K., Fernandez, A., Aida, M., Sedbrook, J., Tasaka, M., Masson, P. and Barton, M.K. (1999). The *PINHEAD/ZWILLE* gene acts pleiotropically in *Arabidopsis* development and has overlapping functions with the *ARGONAUTE1* gene. *Development* **126**, 469-481.

Ma, L., Sun, N., Liu, X., Jiao, Y., Zhao, H. and Deng, X.W. (2005). Organ-specific expression of Arabidopsis genome during development. *Plant Physiology* **138**, 80-91.

Mandel, M.A., Gustafson-Brown, C., Savidge, B. and Yanofsky, M.F. (1992). Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1*. *Nature* **360**, 273-277.

Mansfield, S.G. and Briarty, L.G. (1991). Early embryogenesis in *Arabidopsis thaliana*. II. The developing embryo. *Canadian Journal of Botany* **69**, 461-476.

Mansfield, S.G., Briarty, L.G. and Erni, S. (1991). Early embryogenesis in *Arabidopsis thaliana*. I. The mature embryo sac. *Canadian Journal of Botany* **69**, 447-460.

Marsden, M.P.F. and Meinke, D.W. (1985). Abnormal development of the suspensor in an embryo-lethal mutant of *Arabidopsis thaliana*. *American Journal of Bot*any **72**, 1801-1812.

Martinoia, E., Klein, M., Geisler, M., Bovet, L., Forestier, C., Kokukisaoglu, Ü., Müller-Röber and Schulz, B. (2002). Multifunctionality of plant ABC transporters more than just detoxifiers. *Planta* **214**, 345-355.

Mascarenhas, J. P. and Hamilton, D. A. (1992). Artifacts in localization of GUS activity in anthers of petunia transformed with a CaMV 35S-GUS construct. *The Plant Journal* **2**, 405-408.

Matsunaga, S., Schütze, K., Donnison, I.S., Grant, S.R., Kuroiwa, T. and Kawano, S. (1999). Single pollen typing combined with laser-mediated manipulation. *The Plant Journal* **20**, 371-378.

Mayer, U., Büttner, G. and Jürgens, G. (1993). Apical-basal pattern formation in the *Arabidopsis* embryo: studies on the role of the *GNOM* gene. *Development* **117**, 149-162.

Mayer, U. and Jürgens, G. (1998). Pattern formation in plant embryogenesis: A reassessment. *Seminars in Cell and Developmental Biology* **9**, 187-193.

Mayer, K.F.X., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G. and Laux, T. (1998). Role of *WUSCHEL* in regulating stem cell fate in the *Arabidopsis* shoot meristem. *Cell* **95**, 805-815.

Mayer, U., Torres Ruiz, R.A., Berleth, T., Miséra, S. and Jürgens, G. (1991). Mutations affecting body organization in the *Arabidopsis* embryo. *Nature* **353**, 402-407.

McClintock, B. (1950). The origin and behaviour of mutable loci in maize. *Proceedings of the National Academy of Sciences USA* **36**, 344-355.

McConnell, J.R. and Barton, M.K. (1998). Leaf polarity and meristem formation in *Arabidopsis*. *Development* **125**, 2935-2942.

McConnell, J.R., Emery, J., Eshed, Y., Bao, N., Bowman, J. and Barton, M.K. (2001). Role of *PHABULOSA* and *PHAVOLUTA* in determining radial patterning in shoots. *Nature* **411**, 709-713.

McElver, J., Tzafrir, I., Aux, G., Rogers, R., Ashby, C., Smith, K., Thomas, C., Schetter, A., Zhou, Q., Cushman, M.A., Tossberg, J., Nickle, T., Levin, J.Z., Law, M., Meinke, D. and Patton, D. (2001). Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics* **159**, 1751-1763.

McGall, G., Labadie, j., Brock, P., Wallraff, G., Nguyen, T. and Hinsberg, W. (1996). Light-directed synthesis of high-density oligonucleotide arrays using semiconductor photoresists. *Proceedings of the National Academy of Sciences USA* **93**, 13555-13560.

Meinke, D.W. (1992). A homeotic mutant of *Arabidopsis thaliana* with leafy cotyledons. *Science* **258**, 1647-1650.

Meinke, D.W., Franzmann, L.H., Nickle, T.C. and Yeung, E.C. (1994). *Leafy cotyledon* mutants of *Arabidopsis. The Plant Cell* **6**, 1049-1064.

Meinke, D. W. and Sussex, I. M. (1979). Embryo-lethal mutants of *Arabidopsis thaliana*: a model system for genetic analysis of plant embryo development. *Developmental Biology* **72**, 50-61.

Mewes, H.W., Albermann, K., Heumann, K., Liebl, S. and Pfeiffer, F. (1997). MIPS: a database for protein sequences, homology data and yeast genome information. *Nucleic Acids Research* **25**, 28-30

Meyerowitz, E.M. (1987). *ARABIDOPSIS THALIANA. Annual Review of Genetics* **21**, 93-111.

Miller, M.E. and Chourey, P.S. (1992). The maize invertase-deficient *miniature-1* seed mutation is associated with aberrant pedicel and endosperm development. *The Plant Cell* **4**, 297-305.

Mizukami, Y. and Fischer, R.L. (2000). Plant organ size control: *AINTEGUMENTA* regulates growth and cell numbers during organogenesis. *Proceedings of the National Academy of Sciences USA* **97**, 942-947.

Mòl, R., Matthys-Rochon, E. and Dumas, C. (1994). The kinetics of cytological events during double fertilization in *zea mays L. The Plant Journal* **5**, 197-206.

Morita-Yamamuro, C., Tsutsui, T., Tanaka, A. and Yamaguchi, J. (2004). Knock-out of the plastid ribosomal protein S21 causes impaired photosynthesis and sugar-response during germination and seedling development in *Arabidopsis thaliana*. *Plant Cell Physiology* **45**, 781-788.

Mostov, K.E., Verges, M. and Altschuler, Y. (2000). Membrane traffic in polarized epithelial cells. *Current Opinion in Cell Biology* **12**, 483-490.

Moussian, B., Schoof, H., Haecker, A., Jürgens, G. and Laux, T. (1998). Role of the *ZWILLE* gene in the regulation of central shoot meristem cell fate during *Arabidopsis* embryogenesis. *The EMBO Journal* **17**, 1799-1809.

Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum* **15**, 473-497.

Nakazono, M., Qui, F., Borsuk, L.A. and Schnable, P.A. (2003). Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: Identification of genes expressed differentially in epidermal cells or vascular tissues of maize. *The Plant Cell* **15**, 583-596.

Nevers, P., shepherd, N.A. and Saedler, H. (1986). Plant transposable elements. *Advances in Botanical Research* **12**, 102-103.

Niehrs, C. and Pollet, N. (1999). Synexpression groups in eukaryotes. *Nature* **402**, 483-487.

Nomura, K. and Komamine, A. (1985). Identification and isolation of single cells that produce somatic embryos at a high frequency in a carrot suspension culture. *Plant Physiology* **79**, 988-991.

Nygaard, V., Holden, M., Løland, A., Langaas, M., Myklebost, O. and Hovig, E. (2005). Limitations of mRNA amplification from small-size cell samples. *BMC Genomics* **6**, 147.

Ogas, J., Cheng, J. –C., Renee Sung, Z. and Somerville, C. (1997). Cellular differentiation regulated by gibberellin in the *Arabidopsis thaliana pickle* mutant. *Science* **277**, 91-94.

Ogas, J., Kaufmann, S., Henderson, J. and Somerville, C. (1999). *PICKLE* is a CHD3 chromatin-remodelling factor that regulates the transition from embryonic to vegetative development in *Arabidopsis*. *Proceedings of the National Academy of Sciences USA* **96**, 13839-13844.

Olson, A.R. and Cass, D.D. (1981). Changes in megagametophyte structure in *Papaver nudicaule* following *in vitro* placental pollination. *American Journal of Botany* **68**, 1338-1341.

Ori, N., Eshed, Y., Chuck, G., Bowman, J.L. and Hake, S. (2000). Mechanisms that control *knox* gene expression in the *Arabidopsis* shoot. *Development* **127**, 5523-5532.

Otsuga, D., DeGuzman, B., Prigge, M.J., Drews, G.N. and Clark, S.E. (2001). *REVOLUTA* regulates meristem initiation at lateral positions. *The Plant Journal* **25**, 223-236.

Palme, K. and Galweiler, L. (1999). PIN-pointing the molecular basis of auxin transport. *Current Opinion in Plant Biology* **2**, 375-381.

Pang, P.P., Pruitt, R.E. and Meyerowitz, E.M. (1988). Molecular cloning, genomic organization, expression and evolution of 12S seed storage protein genes of *Arabidopsis thaliana*. *Plant Molecular Biology* **11**, 805-820.

Parcy, F., Valon, C., Kohara, A., Miséra, S. and Giraudat, J. (1997). The *ABSCISIC ACID-INSENSITIVE 3, FUSCA 3* and *LEAFY COTYLEDON 1* loci act in concert to control multiple aspects of *Arabidopsis* seed development. *The Plant Cell* **9**, 1265-1277.

Parcy, F., Valon, C., Raynal, M., Gaubier-Comella, P., Delseny, M. and Giraudat, J. (1994). Regulation of gene expression programs during *Arabidopsis* seed development: Roles of the *ABI3* locus and of endogenous abscisic acid. *The Plant Cell* **6**, 1567-1582.

Parinov, S., Sevugan, M., Ye, D., Yang, W.C., Kumaran, M. and Sundaresan, V. (1999). Analysis of flanking sequences from dissociation insertion lines: a database for reverse genetics in Arabidopsis. *The Plant Cell* **11**, 2263-2270.

Peleman, J., Cottyn, B., Van Camp, W., Van Montagu, M. and Inze, D. (1991). Transient occurence of extrachromosomal DNA of an *Arabidopsis thaliana* transposon-like element, *Tat1*. *Proceedings of the National Academy of Sciences USA* **88**, 3618-3622.

Pereira, A. (1998). Heterologous transposon tagging. *Transgenic Plant Research* **7**, 91-108.

Perlmutter, M.A., Best, C.J.M., Gillespie, J.W., Gathright, Y., González, S., Velasco, A., Linehan, W.M., Emmert-Buck, M.R. and Chuaqui, R.F. (2004). Comparison of snap freezing versus ethanol fixation for gene expression profiling of tissue specimens. *The Journal of Molecular Diagnostics* **6**, 371-377.

Pesaresi, P., Varotto, C., Meurer, J., Jahns, P., Salamini, F. and Leister, D. (2001). Knock-out of the plastid ribosomal protein L11 in *Arabidopsis*: effects on mRNA translation and photosythesis. *The Plant Journal* **27**, 179-189.

Peterson, L.A., Brown, M.R., Carlisle, A.J., Kohn, E.C., Liotta, L.A., Emmert-Buck, M.R. and Krizman, D.B. (1998). An improved method for construction of directionally cloned cDNA libraries from microdissected cells. *Cancer Research* **58**, 5326-5328.

Peyroche, A., Paris, S. and Jackson, C.L. (1996). Nucleotide exchange on ARF mediated by yeast Gea1 protein. *Nature* **384**, 479-484.

Poethig, R.S., Coe, E.H. and Johri, M.M. (1986). Cell lineage patterns in maize embryogenesis: A clonal analysis. *Developmental Biology* **117**, 392-404.

Poirier, G.M.-C. and Erlander, M.G. (1998). Post differential display: Parallel processing of candidates using small amounts of RNA. *Methods: A companion to Method in Enzymology* **16**, 444-452.

Prigge, M.J. and Wagner, D.R. (2001). The Arabidopsis *SERRATE* gene encodes a zinc-finger protein required for normal shoot development. *The Plant Cell* **13**, 1263-1279.

Provart, N.J. and McCourt, P. (2004). Systems approaches to understanding cell signalling and gene regulation. *Current Opinion in Plant Biology* **7**, 605-609.

Przemeck, G.K.H., Mattsson, J., Hardtke, C.S., Sung, Z.R. and Berleth, T. (1996). Studies on the role of the *Arabidopsis* gene *MONOPTEROS* in vascular *Development* and plant cell axialization. *Planta* **200**, 229-237.

Quackenbush, J. (2002). Microarray data normalization and transformation. *Nature Genetics* **32**, 496-501.

Quatrano, R.S. and Shaw, S.L. (1997). Role of the cell wall in the determination of cell polarity and the plane of cell division in *Fucus* embryos. *Trends in Plant Science* **2**, 15-21.

Raghavan, V. (1997) Molecular Embryology of flowering plants (Cambridge, Cambridge University Press).

Ramos, J.A., Zenser, N., Leyser, O., and Callis, J. (2001). Rapid degradation of auxin/indole acetic acid proteins requires conserved amino acids of domain II and is proteasome dependent. *The Plant Cell* **13**, 2349–2360.

Reed, J. W. (2001). Roles and activities of Aux/IAA proteins in *Arabidopsis*. *Trends in Plant Science* **6**, 420–425.

Reidt, W., Wohlfarth, T., Ellerström, M., Czihal, A., Tewes, A., Ezcurra, I., Rask, L. and Bäumlein, H. (2000). Gene regulation during late embryogenesis: the RY motif of maturation-specific gene promoters is a direct target of the *FUS3* gene product. *The Plant Journal* **21** (5), 401-408.

Reinhardt, D. (2005). Phyllotaxis – a new chapter in an old tale about beauty and magic numbers. *Current opinion in plant biology* **8**, 1-7.

Ren, X.-Y., Fiers, M.W.E.J., Stiekema, W.J. and Nap, J.-P. (2005). Local coexpression domains of two to four genes in the genome of Arabidopsis. *Plant Physiology* **138**, 923-934.

Rhee, S.Y., Beavis, W., Berardini, T.Z., Chen, G., Dixon, D., Doyle, A., Garcia-Hernandez, M., Huala, H., Lander, G., Montoya, M., Miller, N., Mueller, L.A., Mundodi, S., Reiser, L., Tacklind, J., Weems, D.C., Wu, Y., Xu, I., Yoo, D., Yoon, J and Zhang, P. (2003). The *Arabidopsis* Information Resource (TAIR): a model organism database providing a centralized, curated gateway to *Arabidopsis* biology, research materials and community. *Nucleic Acids Research* **31**, 224-228.

Ritchie, S. and Gilroy, S. (1998). Gibberellins: regulating genes and germination. *New Phytologist* **140**, 363-383.

Rogg, L.E. and Bartel, B. (2001). Auxin signaling: Derepression through regulated degradation. *Developmental Cell* **1**, 595–604.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B. (2003). An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Molecular Biology* **53**, 247-259.

Russell, L., Larner, V., Kurup, S., Bougourd, S. and Holdsworth, M. (2000). The *Arabidopsis COMATOSE* locus regulates germination potential. *Development* **127**, 3759-3767.

Salunga, R.C., Guo, H., Luo, L., Joy, D.C., Chambers, J.R., Wan, J.S., Jackson, M.R. and Erlander, M.G. (1999). Gene expression analysis via cDNA microarrays of laser capture microdissected cells from fixed tissue. In *DNA Microarrays: A Practical Approach*, M. Schena, ed. (Oxford: Oxford University Press), pp. 121-137.

Sambrook, J., Fritsch, E.F. and Maniatas, T. (1989). Molecular cloning, a laboratory manual, 2nd edition (New York, Cold Spring Harbour Laboratory Press).

Sanchez-Carbayo, M., Saint, F., Lozano, J.J., Viale, A. and Cordon-Cardo, C. (2003). Comparison of gene expression profiles in laser-microdissected, nonembedded, and OCT-embedded tumour samples by oligonucleotide microarray analysis. *Clinical Chemistry* **49**, 2096-2100.

Satou, Y., Takatori, N., Yamada, L., Mochizuki, Y., Hamaguchi, M., Ishikawa, H., Chiba, S., Imai, K., Kano, S., Murakami, S.D., Nakayama, A., Nishino, A., Sasakura, Y., Satoh, G., Shimotori, T., Shin-I, T., Shoguchi, E., Suzuki, M.M., Takada, N., Utsumi, N., Yoshida, N., Saiga, H., Kohara, Y. and Satoh, N. (2001). Gene expression profiles in *Ciona intestinalis* tailbud embryos. *Development* **128**, 2893-2904.

Sawa, S., Watanabe, K., Goto, K., Kanaya, E., Morita, E.H. and Okada, K. (1999). *FILAMENTOUS FLOWER*, a meristem and organ identity gene of *Arabidopsis*, encodes a protein with a zinc finger and HMG-related domains. *Genes and Development* **13**, 1079-1088.

Schell, J., Van Montagu, M., De Beuckleer, M., De Block, M., Depicker, A., De Wilde, M., Engler, G., Genetello, C., Hernalsteens, J.P., Holsters, M., Seurinck, J., Silva, B., Van Vliet, F. and Villarroel, R. (1979). Interactions and DNA transfer between *Agrobacterium tumefaciens*, the Ti-plasmid and the plant host. *Proceedings of the Royal Society of London* B **204**, 251-266.

Scheidl, S.J., Nilsson, S., Kalén, M., Hellström, M., Takemoto, M., Håkansson, J. and Lindahl, P. (2002). mRNA expression profiling of laser microbeam microdissected cells from slender embryonic structures. *American Journal of Pathology* **160**, 801-813.

Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467-470.

Scheres, B., McKhann, H.I. and van den Berg, C. (1996). Roots redefined: anatomical and genetical analysis of root development. *Plant Physiol*ogy **111**, 959-964.

Scheres, B., Wolkenfelt, H., Willemsen, V., Terlouw, M., Lawson, E., Dean, C. and Weisbeek, P. (1994). Embryonic origin of the *Arabidopsis* primary root and root meristem initials. *Development* **120,** 2475-2487.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D. and Lohmann, J.U. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics* **37**, 501-506.

Schneitz, K., Balasubramanian, S. and Schiefthaler, U. (1998). Organogenesis in plants: the molecular and genetic control of ovule development. *Trends in Plant Science* **3,** 468-472.

Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F.X., Jürgens, G. and Laux, T. (2000). The stem cell population of *Arabidopsis* shoot meristems is maintained by a regulatory loop between the *CLAVATA* and *WUSCHEL* genes. *Cell* **100**, 635-644.

Schrick, K., Mayer, U., Horrichs, A., Kuhnt, C., Bellini, C., Dangl, J., Schmidt, J and Jürgens, G. (2000). *FACKEL* is a sterol C-14 reductase required for organised cell expansion in *Arabidopsis* embryogenesis. *Genes and Development* **14**, 1471-1484.

Schultze, A. and Downward, J. (2001). Navigating gene expression using microarrays – a technology review. *Nature Cell Biology* **3**, E190-195

Schulz, R. and Jenson, W.A. (1968). *Capsella* embryogenesis: the egg, zygote and young embryo. *American Journal of Botany* **55**, 807-819.

Schwartz, B.W., Yeung, E.C. and Meinke, D.W. (1994). Disruption of morphogenesis and transformation of the suspensor in abnormal *sus*pensor mutants of *Arabidopsis. Development* **120**, 3235-3245.

Semiarti, E., Ueno, Y., Tsukaya, H., Iwakawa, H., Machida, C. and Machida, Y. (2001). The *ASYMMETRIC LEAVES2* gene of *Arabidopsis thaliana* regulates formation of a symmetric lamina, establishment of venation and repression of meristem-related homeobox genes in leaves. *Development* **128**, 1771-1783.

Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A. (2002). A high-throughput Arabidopsis reverse genetics system. *The Plant Cell* **14**, 2985-2994.

Shaw, C.H. (1995). Introduction of cloning plasmids into *Agrobacterium tumefaciens.* In Plant Gene Transfer and Expression Protocols, Methods in Molecular Biology, vol. 49. Jones, H. ed(Humana Press: Totowa, N.J.).

Shaw, S.L. and Quatrano, R.S. (1996). The role of targeted secretion in the establishment of cell polarity and the orientation of the division plane in *Fucus* zygotes. *Development* **122**, 2623-2630.

Sheen, J., Hwang, S., Niwa, Y., Kobayashi, H. and Galbraith, D.W. (1995). Green-fluorescent protein as a new vital marker in plant cells. *The Plant Journal.* **8**, 777-784.

Sheridan, W.F. and Clarke, J.K. (1987). Maize embryogeny: a promising experimental system. *Trends in Genetics* **3**, 3-6.

Shevell, D.E., Leu, W.-M., Gillmor, C.S., Xia, G., Feldmann, K.A. and Chua, N.-H. (1994). *EMB30* is essential for normal cell division, cell expansion, and cell adhesion in *Arabidopsis* and encodes a protein that has similarity to Sec7. *Cell* **77**, 1051-1062.

Shiu, S.-H. and Bleecker, A.B. (2003). Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. *Plant Physiology* **132**, 530-543.

Showalter, A.M. (2001). Introduction: plant cell wall proteins. *Cellular and Molecular Life Sciences* **58**, 1361-1362.

Siegfried, K.R., Eshed, Y., Baum, S.F., Otsuga, D., Drews, G.N. and Bowman, J.L. (1999). Members of the *YABBY* gene family specify abaxial cell fate in *Arabidopsis. Development* **126**, 4117-4128.

Simone, N.L., Bonner, R.F., Gillespie, J.W., Emmert-Buck, M.R., Liotta, L.A. (1998). Laser-capture microdissection: opening the microscopic frontier to molecular analysis. *Trends in Genetics* **14**, 272-276.

Sinha, N.R., Williams, R.E. and Hake, S. (1993). Overexpression of the maize homeobox gene, *KNOTTED-1*, causes a switch from determinate to indeterminate cell fates. *Genes and Development.* **7**, 787-795.

Sivamarakrishna, D. (1978). Size relationship of apical and basal cell in two-celled embryos in angiosperms. *Canadian Journal of Botany* **56**, 1434-1438.

Snow, M. and Snow, R. (1959). The dorsoventrality of leaf primordia. *New Phytologist* **5**, 286-300.

Souer, E., Van Houwelingen, A., Kloos, D., Mol, J. and Koes, R. (1996). The *NO APICAL MERISTEM* gene of petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. *Cell* **85**, 159-170.

Souter, M. and Lindsey, K. (2000). Polarity and signalling in plant embryogenesis. *Journal of Experimental Botany* **51**, 971-983.

Souter, M. and Lindsey, K. (2001). Orchestrating morphogenesis: The importance of signalling in embryogenesis. *Phytomorphology Golden Jubilee Issue*, 305-325.

Souter, M., Topping, J., Pullen, M., Friml, J., Palme, K., Hackett, R., Grierson, D. and Lindsey, K. (2002). *hydra* mutants are defective in sterol profiles and auxin and ethylene signalling. *The Plant Cell* **14**, 1017-1031.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503-517.

Spellman, P.T., Sherlock, G., Zhang, M.O., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D. and Futcher, B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Molecular Biology of the Cell* **9**, 3273-3297.

Springer, P.S., McCombie, W.R., Sundaresan, V. and Martienssen, R.A. (1995). Gene trap tagging of *PROLIFERA*, an essential MCM2-3-5-like gene in *Arabidopsis*. *Science* **268**, 877-880.

Srinivasan, M., Sedmak, D. and Jewell, S. (2002). Effect of fixatives and tissue processing on the content and integrity of nucleic acids. *American Journal of Pathology* **161**, 1961-1971.

St Johnston, D. and Nüsslein-Volhard, C. (1992). The origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68**, 201-219.

Stachel, S.E. and Nester, E.W. (1986). The genetic and transcriptional organization of the *vir* region of the A6 Ti plasmid of *Agrobacterium tumefacians*. *The EMBO Journal* **5**, 1445-1454.

Standaert, D.G. (2005). Applications of laser capture microdissection in the study of neurodegenerative disease. *Archives of Neurol*ogy **62**, 203-205.

Steinmann, T., Geldner, N., Grebe, M., Mangold, S., Jackson, C.L., Paris, S., Gälweiler, L., Palme, K. and Jürgens, G. (1999). Coordinated polar localization of auxin efflux carrier *PIN1* by *GNOM ARF GEF*. *Science* **286**, 316-318.

Steeves, T.A. and Sussex, I.M. (1989). Patterns in Plant Development, 2edn. (Cambridge: Cambridge University Press).

Sterk, P., Booij, H., Schellekens, G.A., Van Kammen, A. and de Vries, S.C. (1991). Cell-specific expression of the carrot EP2 lipid transfer protein gene. *The Plant Cell* **3**, 907-921.

Stolc, V., Gauhar, Z., Mason, C., Halasz, G., van Batenburg, M.F., Rifkin, S.A., Hua, S., Herreman, T., Tongprasit, W., Barbano, P.E., Bussemaker, H.J. and White, K.P. (2004). A gene expression map for the euchromatic genome of *Drosophila melanogaster*. *Science* **306**, 655-660.

Stomp, A.M. (1990). Use of X-Gluc for histochemical localisation of glucuronidase. In *Editorial Comments*. Cleveland: United State Biochemical, p. 5.

Stone, S.L., Kwong, L.W., Yee, K.M., Pelletier, J., Lepiniec, L., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2001). *LEAFY COTYLEDON2* encodes a B3 domain transcription factor that induces embryo development. *Proceedings of the National Academy of Sciences USA* **98**, 11806-11811.

Stoughton, R. (2005). Applications of DNA microarrays in biology. *Annual Review of Biochemistry* **74**, 53-82.

Sussex, I.M. (1954). Experiments on the cause of dorsoventrality in leaves. *Nature* **174**, 351-352.

Swamy, B.G.L. and Krishnamurthy, K.V. (1977). Certain conceptual aspects of meristems. II. Epiphysis and shoot apex. *Phytomorphology* **27**, 1-8.

Takada, S., Hibara, K. –I., Ishida, T. and Tasaka, M. (2001). The *CUP-SHAPED COTYLEDON1* gene of *Arabidopsis* regulates shoot apical meristem formation. *Development* **128**, 1127-1135.

Talbert, P.B., Adler, H.T., Parks, D.W. and Comai, L. (1995). The *REVOLUTA* gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of *Arabidopsis thaliana*. *Development* **121**, 2723-2735.

The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815.

Thoma, S., Hecht, U., Kippers, A., Botella, J., de Vries, S. and Somerville, C. (1994). Tissue-specific expression of a gene encoding a cell wall-localized lipid transfer protein from *Arabidopsis*. *Plant physiology* **105**, 35-45.

Thomas, C. (1996). PCR Techniques. In Plant Gene Isolation, G. D. Foster and D. Twell, ed (Chichester, John Wiley & Sons), pp 331-368.

Timmermans, M.C., Hudson, A., Becraft, P.W. and Nelson, T. (1999). ROUGH SHEATH2: a Myb protein that represses Knox homeobox genes in maize lateral organ primordia. *Science* **284**, 151-153.

Tiwari, S.B., Wang, X.-J., Hagen, G., and Guilfoyle, T.J. (2001). AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin. *The Plant Cell* **13**, 2809–2822.

Tomancak, P., Beaton, A., Weiszmann, R., Kwan, E., Shu, S., Lewis, S.E., Richards, S., Ashburner, M., Hartenstein, V., Celniker, S.E. and Rubin, G.M. (2002). Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biology* **3**, RESEARCH0088.1-RESEARCH0088.14.

Topping, J.F., Agyeman, F., Henricot, B. and Lindsey, K. (1994). Identification of molecular markers of embryogenesis in *Arabidopsis thaliana* by promoter trapping. *The Plant Journal* **5**, 895-903.

Topping, J.F. and Lindsey, K. (1995). Insertional mutagenesis and promoter trapping in plants for the isolation of genes and the study of development. *Transgenic Research* **4**, 291-305.

Topping, J.F. and Lindsey, K. (1997). Promoter trap markers differentiate structural and positional components of polar development in *Arabidopsis*. *The Plant Cell* **9**, 1713-1725.

Topping, J.F., Wei, W. and Lindsey, K. (1991). Functional tagging of regulatory elements in the plant genome. *Development* **112**, 1009-1019.

Torres Ruiz, R.A. (2004). Polarity in *Arabidopsis* embryogenesis. In Polarity of Plants, K. Lindsey, ed(Oxford: Blackwell Publishing), pp 157-191.

Torres Ruiz, R.A. and Jürgens, G. (1994). Mutations in the *FASS* gene uncouple pattern-formation and morphogenesis in *Arabidopsis* development. *Development* **120**, 2967-2978.

Torres Ruiz, R.A., Lohner, A. and Jürgens, G. (1996). The *GURKE* gene is required for normal organization of the apical region in the *Arabidopsis* embryo. *The Plant Journal* **10**, 1005-1016.

Trotochaud, A.E., Jeong, S. and Clark, S.E. (2000). CLAVATA3, a multimeric ligand for the CLAVATA1 receptor-kinase. *Science* **289**, 613-617.

Tsay, Y.-F., Frank, M. J., Page, T., Dean, C. and Crawford, N. M. (1993). Identification of a mobile endogenous transposon in *Arabidopsis thaliana*. *Science* **260**, 342-344.

Tsiantis, M., Schneeberger, R., Golz, J.F., Freeling, M. and Langdale, J.A. (1999). The maize *rough sheath2* gene and leaf development programs in monocot and dicot plants. *Science* **284**, 154-156.

Tsuchiya, Y., Nambara, E., Naito, S. and McCourt, P (2004). The *FUS3* transcription factor functions through the epidermal regulator *TTG1* during embryogenesis in *Arabidopsis*. *The Plant Journal* **37**, 73-81.

Tykarska, T. (1976). Rape embryogenesis. I. The proembryo development. *Acta Societatis Botanicorum Poloniae* **45**, 3-15.

Tykarska, T. (1979). Rape embryogenesis. II. Development of the embryo proper. *Acta Societatis Botanicorum Poloniae* **48**, 391-421.

Tykarska, T. (1987). Rape embryogenesis. V. Accumulation of lipid bodies. *Acta Societatis Botanicorum Poloniae* **56**, 573-584.

Tzafrir, I., McElver, J.A., Liu, C.M., Yang, L.J., Wu, J.Q., Martinez, A., Patton, D.A. and Meinke, D. (2002). Diversity of TITAN functions in *Arabidopsis* seed development. *Plant Physiology* **128**, 38-51.

Tzafrir, I., Pena-Muralla, R., Dickermann, A., Berg, M., Rogers, R., Hutchens, S., Sweeney, T.C., McElver, J., Aux, G., Patton, D. and Meinke, D. (2004). Identification of genes required for embryo development in Arabidopsis. *Plant Physiology* **135**, 1206-1220.

Ulmasov, T., Hagen, G. and Guilfoyle, T.J. (1997). ARF1, a transcription factor that binds to auxin response elements. *Science* **276**, 1865-1868.

Ulmasov, T., Hagen, G. and Guilfoyle, T.J. (1999). Activation and repression of transcription by auxin-response factors. *Proceedings of the National Academy of Science USA* **96**, 5844-5849.

van den Berg, C., Willemsen, V., Hage, W., Weisbeek, P. and Scheres, B. (1997). Short-range control of cell differentiation in the *Arabidopsis* root meristem. *Nature* **390**, 287-289.

Van Gelder, R.N., von Zastrow, M.E., Yool, A., Dement, W.C., Barchas, J.D. and Eberwine, J.H. (1990). Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proceedings of the National Academy of Sciences USA* **87**, 1663-1667.

van Lammeren, A.A.M. (1981). Early events during embryogenesis in *Zea mays* L. *Acta Societatis Botanicorum Poloniae* **50**, 289-290.

Van Sluys, M.A., Tempé, J and Fedoroff, N. (1987). Studies on the introduction and mobility of the maize Activator element in Arabidopsis thaliana and Daucus carota. *The EMBO Journal* **6**, 3881-3889.

Vernon, D.M. and Meinke, D.W. (1994). Embryogenic transformation of the suspensor in *twin*, a polyembryonic mutant of *Arabidopsis*. *Developmental Biology* **165**, 566-573.

Vittorioso, P., Cowling, R., Faure, J. –D., Caboche, M. and Bellini, C. (1998). Mutation in the *Arabidopsis PASTICCINO1* gene, which encodes a new FK506-binding protein-like protein, has a dramatic effect on plant development. *Molecular and Cellular Biology* **18**, 3034-3043.

Vroemen, C.W., Langeveld, S., Mayer, U., Ripper, G., Jürgens, G., Van Kammen, A. and de Vries, S.C. (1996). Pattern formation in the Arabidopsis embryo revealed by position-specific lipid transfer protein gene expression. *The Plant Cell* **8**, 783-791.

Vroemen, C.W., Mordhorst, A.P., Albrecht, C., Kwaaitaal, M.A.C.J. and de Vries, S.C. (2003). The *CUP-SHAPED COTYLEDON3* gene is required for boundary and shoot meristem formation in Arabidopsis. *The Plant Cell* **15**, 1563-1577.

Walbot, V. (1978). Control mechanisms for plant embryogeny. In Dormancy and developmental arrest, M.E. Clutter, ed(New York: Academic Press), pp 113-166.

Walbot, V. (2000). A green chapter in the book of life. *Nature* **408**, 794-795.

Waites, R., Selvadurai, H.R.N., Oliver, I.R. and Hudson, A. (1998). The *PHANTASTICA* gene encodes a MYB transcription factor involved in growth and dorsoventrality of lateral organs in *Antirrhinum*. *Cell* **93**, 779-789.

Waites, R. and Simon, R. (2000). Signalling cell fate in plant meristems: Three clubs on one tousle. *Cell* **103**, 835-838.

Weijers, D., Sauer, M., Meurette, O., Friml, J., Ljung, K., Sandberg, G., Hooykaas, P. and Offringa, R. (2005). Maintenance of embryonic auxin distribution for apical-basal patterning by PIN-FORMED-dependent auxin transport in *Arabidopsis*. *The Plant Cell* **17**, 2517-2526.

Weir, I., Lu, J., Cook, H., Causier, B., Schwartz-Sommer, Z. and Davies, B. (2004). *CUPULIFORMIS* establishes lateral organ boundaries in *Antirrhinum*. *Development* **131**, 915-922.

Wellmer, F., Riechmann, J.L., Alves-Ferreira, M. and Meyerowitz, E.M. (2004). Genome-wide analysis of spatial gene expression in Arabidopsis flowers. *The Plant Cell* **16**, 1314-1326.

West, M.A.L. and Harada, J.J. (1993). Embryogenesis in higher plants: An overview. *The Plant Cell* **5**, 1361-1369.

West, M.A.L., Yee, K.M., Danao, J., Zimmerman, J.L., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (1994). *LEAFY COTYLEDON1* is an essential regulator of late embryogenesis and cotyledon identity in Arabidopsis. *The Plant Cell* **6**, 1731-1745.

White, K.P., Rifkin, S.A., Hurban, P., Hogness, D.S. (1999). Microarray analysis of *Drosophila* development during metamorphosis. *Science* **286**, 2179-2184.

Whittington, A.T., Vugrek, O., Wei, K.J., Hasenbein, N.G., Sugimoto, K., Rashbrooke, M.C. and Wasteneys, G.O. (2001). MOR1 is essential for organizing cortical microtubules in plants. *Nature* **411**, 610-613.

Willemsen, V., Wolkenfelt, H., de Vrieze, G., Weisbeek, P. and Scheres, B. (1998). The *HOBBIT* gene is required for formation of the root meristem in the *Arabidopsis* embryo. *Development* **125**, 521-531.

Woodrick, R., Martin, P.R., Birman, I. and Picket, F.B. (2000). The Arabidopsis embryonic shoot fate map. *Development* **127**, 813-820.

Worley, C.K., Zenser, N., Ramos, J., Rouse, D., Leyser, O., Theologis, A., and Callis, J. (2000). Degradation of Aux/IAA proteins is essential for normal auxin signalling. *The Plant Journal* **21**, 553–562.

Wortman, J.R., Haas, B.J., Hannick, L.I., Smith, R.K.jr., Maiti, R., Ronning, C.M., Chan, A.P., Yu, C., Ayele, M., Whitelaw, C.A., White, O.R. and Town, C.D. (2003). Annotation of the Arabidopsis genome. *Plant Physiology* **132**, 461-468.

Wu, L.F., Hughes, T.R., Davierwala, A.P., Robinson, M.D., Stoughton, R. and Altschuler, S.J. (2002). Large-scale prediction of *Saccharomyces cerevisiae* gene function using overlapping transcriptional clusters. *Nature Genetics* **31**, 255-265.

Wu, Z., Irizarry, R.A., Gentleman, R., Murillo, F.M. and Spencer, F. (2004). A model based background adjustment for oligonucleotide expression arrays. *In Dept. of Biostatistics Working Papers. Working Paper 1.*(John Hopkins University).

Yadegari, R., de Paiva, G.R., Laux, T., Koltunow, A.M., Apuya, N., Zimmerman, J.L., Fischer, R.L., Harada, J.J. and Goldberg, R.B. (1994). Cell differentiation and morphogenesis are uncoupled in Arabidopsis *raspberry* embryos. *The Plant Cell* **6**, 1713-1729.

Yamada, K., Lim, J., Dale, J.M., Chen, H., Shinn, P., Palm, C.J., Southwick, A.M., Wu, H.C., Kim, C., Nguyen, M., Pham, P., Cheuk, R., Karlin-Newmann, G., Liu, S.X., Lam, B., Sakano, H., Wu, T., Yu, G., Miranda, M., Quach, H.L., Tripp, M., Chang, C.H., Lee, J.M., Toriumi, M., Chan, M.M.H., Tang, C.C., Onodera, C.S., Deng, J.M., Akiyama, K., Ansari, Y., Arakawa, T., Banh, J., Banno, F., Bowser, L., Brooks, S., Carninci, P., Chao, Q., Choy, N., Enju, A., Goldsmith, A.D., Gurjal, M., Hansen, N.F., Hayashizaki, Y., Johnson-Hopson, C., Hsuan, V.W., Lida, K., Karnes, M., Khan, S., Koesema, E., Ishida, J., Jiang, P.X., Jones, T., Kawai, J., Kamiya, A., Meyers, C., Nakajima, M., Narusaka, M., Seki, M., Vaysberg, M., Wallender, E.K., Wong, C., Yamamura, Y., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R. (2003). Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* **302**, 842-846.

Yeung, E.C. and Meinke, D.W. (1993). Embryogenesis in angiosperms: development of the suspensor. *The Plant Cell* **5**, 1371-1381.

Yeung, E.C. and Sussex, I.M. (1979). Embryogeny of *Phaseolus coccineus*: The suspensor and the growth of the embryo-proper *in vitro*. *Z. Pflanzenphysiol.* **91**, 423-433.

Yeung, K.Y., Medvedovic, M. and Bumgarner, R.E. (2004). From co-expression to co-regulation: how many microarray experiments do we need? *Genome Biology* **5**, R48.

Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Chen, Y., Tong, W., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Huang, X., Li, W., Li, J., Liu, Z., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, J., Zhu, L., Yuan, L. and Yang, H. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**, 79-92.

Yu, L.P., Simon, E.J., Trotochaud, A.E. and Clark, S.E. (2000). *POLTERGEIST* functions to regulate meristem development downstream of the *CLAVATA* loci. *Development* **127**, 1661-1670.

Zambryski, P., Joos, H., Genetello, C., Leemans, J., Van Montagu, M. and Schell, J. (1983). Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regenerative capacity. *The EMBO Journal* **2**, 2143-2150.

Zenser, N., Ellsmore, A., Leasure, C., and Callis, J. (2001). Auxin modulates the degradation rate of Aux/IAA proteins. *Proceedings of the National Academy of Sciences USA* **98**, 11795–11800.

Zhang, J.Z. and Somerville, C.R. (1997). Suspensor-derived polyembryony caused by altered expression of valyl-tRNA synthetase in the *twn2* mutant of *Arabidopsis*. *Proceedings of the National Academy of Sciences USA* **94**, 7349-7355.

Zhu, T. and Wang, X. (2000). Large-scale profiling of the *Arabidopsis* transcriptome. *Plant Physiology* **124**, 1472-1476.

Zimmerman, J.L. and Goldberg, R.B. (1977). DNA sequence organization in the genome of *Nicotiana tobacum*. *Chromosome* **59**, 227-252.

Zupan, J.R. and Zambryski, P. (1995). Transfer of T-DNA from *Agrobacterium* to the plant cell. *Plant Physiology* **107**, 1041-1047.

Appendices

Appendix 1.

Plant growth media

Murashige and Skoog basal medium (Murashige and Skoog, 1962)

Component	mg/litre
ammonium nitrate	1650.0
boric acid	6.2
calcium chloride (anhydrous)	332.2
cobalt chloride.$6H_2O$	0.025
cupric sulphate.$5H_2O$	0.025
EDTA (disodium)	37.26
ferrous sulphate.$7H_2O$	27.8
magnesium sulphate	180.7
manganese sulphate.H_2O	16.9
molybdic acid (sodium salt).$2H_2O$	0.25
potassium iodide	0.83
potassium nitrate	1900.0
potassium phosphate monobasic	170.0
zinc sulphate.$7H_2O$	8.6
glycine (free base)	2.0
myo-inositol	100.0
nicotinic acid (free acid)	0.5
pyridoxine-HCl	0.5
thiamine-HCl	0.1

1/2MS10

The seed germination medium 1/2 MS10 was prepared by dissolving MS basal medium (half concentration) and 1% w/v sucrose in distilled water with the pH altered to 5.8 with 1M KOH. The medium was solidified with 8g/l bactoagar and sterilised by autoclaving at 121°C for 20 minutes.

Appendix 2.

Primers

1. RT-PCR primers

Table 1 RT-PCR primers (Section 3.5)

Oligo Name	Oligo Details (5'-3')
ANT Forward	CAAGCACGGATTGGTAGAGTCG
ANT Reverse	CATTAGCGTTTGATGTCCAAGG
PINOID Forward	CTCTCTCCGTCATAGACAACC
PINOID Reverse	GCATTACCATGTGATCCACCTG
MP Forward	GATGTCCAGTCGCAGATCACATC
MP Reverse	CTCATCTGCTGGACCTCAGTTGG
BDL Forward	GCGTGGTGTGTCAGAATTGGAGG
BDL Reverse	CTGCCTATACCAACTCCATCC
AtPIN4 Forward	CATTGCTTGTGGGAACTCTGTC
AtPIN4 Reverse	CTATTCCTTGAGGCAACGCAGC
ACT3 Forward	CCATCTTGGCCTCCCTCAGTACC
ACT3 Reverse	CACTCAAACACAGTTGTGTCATCC

Table2 RT-PCR primers (Section 5.6.3)

Oligo Name	Oligo Details (5'-3')
At2g23170 Forward	CCATCACAGAGTTCCTCACAAGC
At2g23170 Reverse	GGAATCCAATGGCACGGAGGAGACC
At5g48490 Forward	GACAAGCAAGAAGGTGGC
At5g48490 Reverse	GTTGGGGCGTTGGTTAGGTCACAC
At1g68780 Forward	CGTGTGGTGTTTGGTGGTGGATTGG
At1g68780 Reverse	CAGAAGGGAGTTCACCTATCAG
At2g03870 Forward	GGTGTGCAAGTTAAGCTCACTGG
At2g03870 Reverse	GACAGCCTCTGCAGTAACGAACG
At5g47500 Forward	GAGAGAAAGTGGTGGTTCCAGC
At5g47500 Reverse	GCTATCGACCCGAATCTTGACG
At3g63040 Forward	GTGCAACAGTGTCAAGTGCAGTAG
At3g63040 Reverse	CTCTCATGGCAAATCCTCATGCACC
At2g23510 Forward	GGTGTGCCGTTCACGGTAGCTACTGC
At2g23510 Reverse	CCGTTGGTAAGTAAGGCGAAGCAGC
At3g13520 Forward	GGAGTCAATGAAGATGAAGCTC
At3g13520 Reverse	CCTGATGCCAAAGCAGCAACAG
At1g20450 Forward	CCAGAGCAGGAGACCCCTAAGGTTGC
At1g20450 Reverse	CCTGGTTTCTCTCCGAGTGGAAAC
At3g54260 Forward	GGATCCCTAAGGCTTGTTCACTACC
At3g54260 Reverse	CCATGTGGTTAGTCCTCTCTC
ACT3 Forward	CCATCTTGGCCTCCCTCAGTACC
ACT3 Reverse	CACTCAAACACAGTTGTGTCATCC

2. Promoter::GUS construct primers

Table 3 Promoter::GUS construct primers (Section 4.5.1)

Oligo Name	Oligo Details (5'-3')
At1g63900/A Forward Outer	CGTTGCGCTTCTCACAATCTTTTG
At1g63900/A Reverse Outer	GCTTCGTAATTAGGAACACACC
At1g63900/A Forward Inner	TTGGATCCCCCGTTGCGGTTAAACTGTAGTCG
At1g63900/A Reverse Inner	TTGGATCCGGAGGCATATTTGTACAACC
At1g63900/A Forward Internal 3	CCAAAACTGTTTGGGATAGGACG
At1g78160/B Forward Outer	GGTGTTTGGTGGTCAAGTCACCG
At1g78160/B Reverse Outer	CGAAGAGACCGAAGAAGCTTCACG
At1g78160/B Forward Inner	TTGGATCCCCGCAGGTTAAACTTACTTCC
At1g78160/B Reverse Inner	TTGGATCCCTCTGTAACAACAACACTAACCC
At1g78160/B Forward Internal 3	CTGTACGTGACATCGCTATTGC
At5g14610/C Forward Outer	CCAACTGTCATAGGCATATAAGTCC
At5g14610/C Reverse Outer	CCTCAGGAGCGTAACGAATTGCAG
At5g14610/C Forward Inner	TTGGATCCCCGGTTTTTGCTCATGATGAGTTTG
At5g14610/C Reverse Inner	TTGGATCCCCTGGAAGAACAACATGTCGAAACC
At5g14610/C Forward Internal 3	GGAATTGAGAGGGCCCTCG
At5g43040/D Forward Outer	GCGATCAACTCTTGAAATTAGGG
At5g43040/D Reverse Outer	CCACCATCAGCACTCTCATGTG
At5g43040/D Forward Inner	TTGGATCCCTATGAAAATCAATACAGCCATGGG
At5g43040/D Reverse Inner	TTGGATCCCAAAAAACTTTGGGGGTGAGG
At5g43040/D Forward Internal 3	GAAGGTCATATTTTCTGCAGC
At5g50810/E Forward Outer	GCGGATTCTGCTTTTCCTTTAG
At5g50810/E Reverse Outer	GCAATTCCGGGTTGTTTGCC
At5g50810/E Forward Inner	TTGGATCCCTGCAAATTCAGAGATCGTAG
At5g50810/E Reverse Inner	TTGGATCCGTCTGGTGAGCTGAGAGCGTCTCTGC
At5g50810/E Forward Internal 3	CCATACATACACTTGAACG
At2g45050/F Forward Outer	CGTGTTCTTGCTGTTTACTTTGG
At2g45050/F Reverse Outer	GGTGAAGATAAGCCATAGACGTCC
At2g45050/F Forward Inner	TTGGATCCGTCACGTCACGTGATTCAAGAAAC
At2g45050/F Reverse Inner	TTGGATCCGTTCTCTCTTGCAGACGAAGACTC
At2g45050/F Forward Internal 3	GGATATAACTCTATTTGCTGGTTGG

At2g25420/Q Forward	TTGTCGACGACCTAAACCATAGCCGTAAATCGG
At2g25420/Q Reverse	TTGTCGACCAAACAAGCATTGAGGTCCAGAG
At2g31510/R Forward	TTGGATCCCATGGAGTGCACGTTTCCTCTCG
At2g31510/R Reverse	TTGGATCCGATCAGAGAAAACGAAATGGC
At3g60860/S Forward	TTGTCGACGATCCGACGGCTGACATACTCG
At3g60860/S Reverse	TTGTCGACGCGAGCTCAGCGACTCACAG
At3g60860/S Forward Internal 3	CTCGAAATGGTATGACTCTCC
At5g45600/T Forward	GTAGTGATGATACTCAAGCACACC
At5g45600/T Reverse	CTCGGCTTAACTTCAACAGATCTGCTTC
GUS NESTED1	GCGATCCAGACTGAATGC

3. SALK primers

Table 4 SALK line primers (Section 4.6.1)

Oligo Name	Oligo Details (5'-3')
SALK_012862 Forward	CCACATTATTCAACACTTGTCTC
SALK_012862 Reverse	GTCGATTCACATGGAAACTTGAGC
SALK_033446 Forward	GCAATGGATTCTCTCCGCCAG
SALK_033446 Reverse	GCTATAGAACTGAGACTGATATCC
SALK_063571 Forward	GTGACTTCAGAGTCTGATGATTCC
SALK_063571 Reverse	CCAGAAACCAAGGAACCTCTTTGC
SALK_068359 Forward	CAAGTATTTTCACTTACCGCACACG
SALK_068359 Reverse	GCCAGCCATACTTGTCTTTTGAGTG
SALK_068825 Forward	CCTTCTTTACGTGGCAAACTTG
SALK_068825 Reverse	CAAGAGATCCAGATGGGGTTGC
SALK_106430 Forward	GCTTGAACCTATCCAACAAAGAACC
SALK_106430 Reverse	CACAAGTGGGCAGTATATGTTCG
SALK_116644 Forward	GCTTCTGCAATTCGTTACGCTCC
SALK_116644 Reverse	CCTGTCACCACATAGGTAGAGC
SALK_131239 Forward	GCTGTAGATGATATGAGGCAAGG
SALK_131239 Reverse	CCCAGTTTGTAAAAGCTCCCAG
SALK_132004 Forward	GGCCAATTGAAAAGAATCAGGGC
SALK_132004 Reverse	CAGCCATAGCCTTTTGCCTCG
SALK_135981 Forward	CATTTTTCTGCCTAGAACACAC
SALK_135981 Reverse	GTAACAACAACACTAACCCAACG
SALK_144520 Forward	CCCACATTCTCTTACAATTCACC
SALK_144520 Reverse	CGTACACTCGGCCCAGATCATGATCC
SALK_508845 Forward	GAGGTTTGATTTTCCAACGTTCAC
SALK_508845 Reverse	GCTGATTTTCCGGCGAATCCATTAGG
pROK2.LB1	TGGTTCACGTAGTGGGCCATCG

Appendix 3

Promoter::GUS constructs

Figure 1 Cloning of GUS Construct A (At1g63900)

A. **Promoter PCR – Outer Reaction**

Lane 2: ~2.5 kb promoter PCR fragment (no restriction sites).

B. **Promoter PCR – Inner Reaction (addition of restriction sites)**

Lane 2: ~2.5 kb promoter PCR fragment (BamH I restriction sites).

C. **Colony PCR for promoter-TOPO colonies**

Asterisked lanes: Putative positive colonies (~2.5 kb product).

D. **Promoter fragment digested with BamH I (prior to purification)**

Asterisked lane: ~2.5 kb promoter fragment and ~4 kb linearised TOPO vector.

E. **pΔGUS-CIRCE digested with BamH I**

Lanes 2-4: BamH I digested pΔGUS-CIRCE vector.

Lane 5 (asterisked): Un-digested pΔGUS-CIRCE vector.

F. **Colony PCR for Promoter-pΔGUS-CIRCE colonies**

Lanes 3-5, 8-12: Putative positive colonies (~800 bp product).

G. **Colony PCR for successfully transformed agrobacterium colonies**

Asterisked lanes: Putative positive colonies (~800 bp product).

In all cases, lane 1 is Hyperladder I (Bioline), 10 kb – 100 bp.

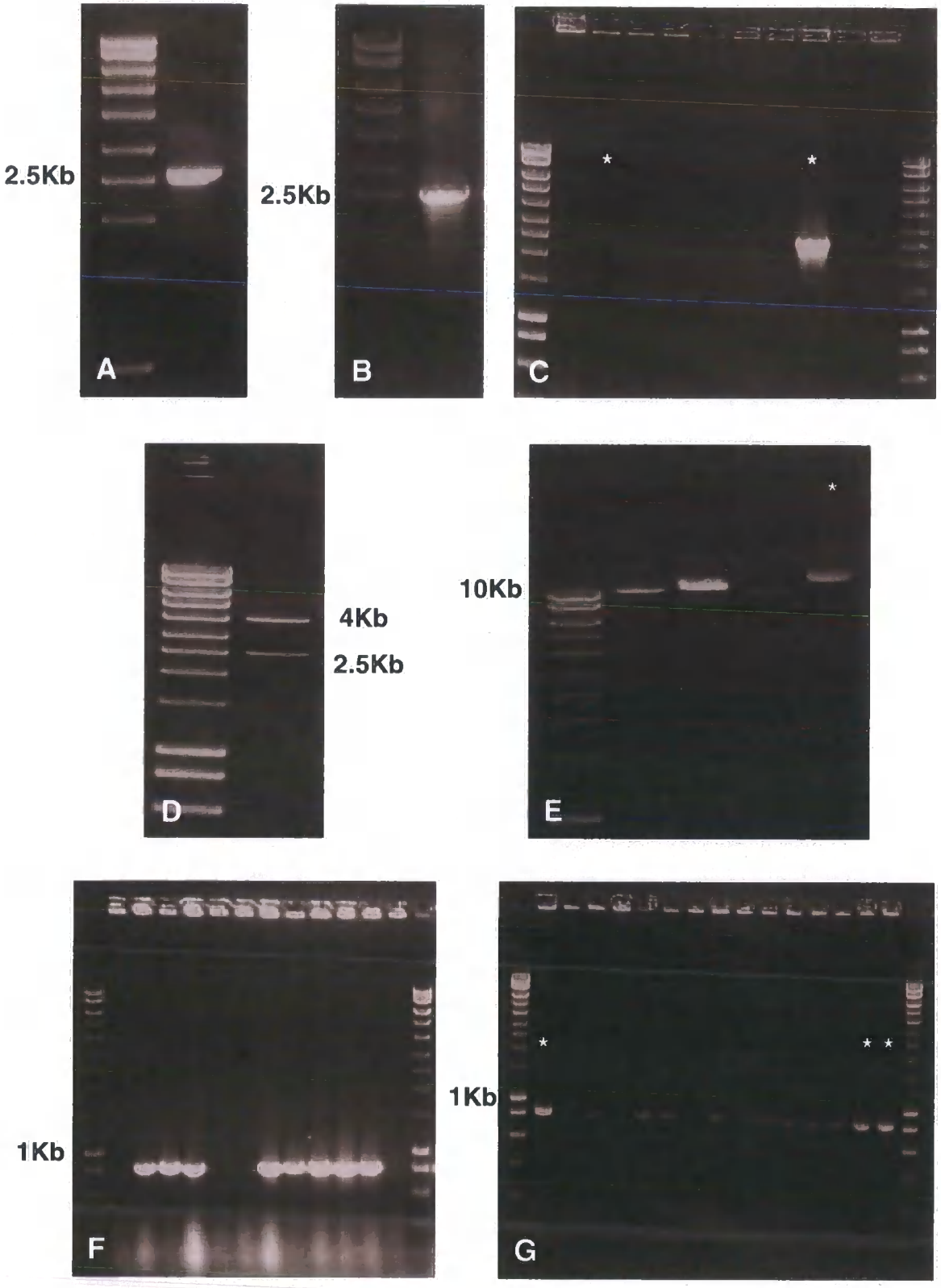

2.5Kb
A
2.5Kb
B
C
4Kb
2.5Kb
D
10Kb
E
1Kb
F
1Kb
G

Figure 2 Cloning of GUS Construct B (At1g78160)

A. Promoter PCR – Outer Reaction

Asterisked lane: ~2.5 kb promoter PCR fragment (no restriction sites).

B. Promoter PCR – Inner Reaction (addition of restriction sites)

Lane 2: ~2.5 kb promoter PCR fragment (BamH I restriction sites).

C. Colony PCR for promoter-TOPO colonies

Asterisked lane: Putative positive colony (~2.5 kb product).

D. Promoter fragment digested with BamH I (prior to purification)

Asterisked lane: ~2.5 kb promoter fragment and ~4 kb linearised TOPO vector.

E. pΔGUS-CIRCE digested with BamH I

Lanes 2-4: BamH I digested pΔGUS-CIRCE vector.

Lane 5 (asterisked): Un-digested pΔGUS-CIRCE vector.

F. Colony PCR for Promoter-pΔGUS-CIRCE colonies

Asterisked lanes: Putative positive colonies (~1 kb product).

G. Colony PCR for successfully transformed agrobacterium colonies

Lanes: 2-3, 5, 7-12, 14-15: Putative positive colonies (~1 kb product).

In all cases, lane 1 is Hyperladder I (Bioline), 10 kb – 100 bp.

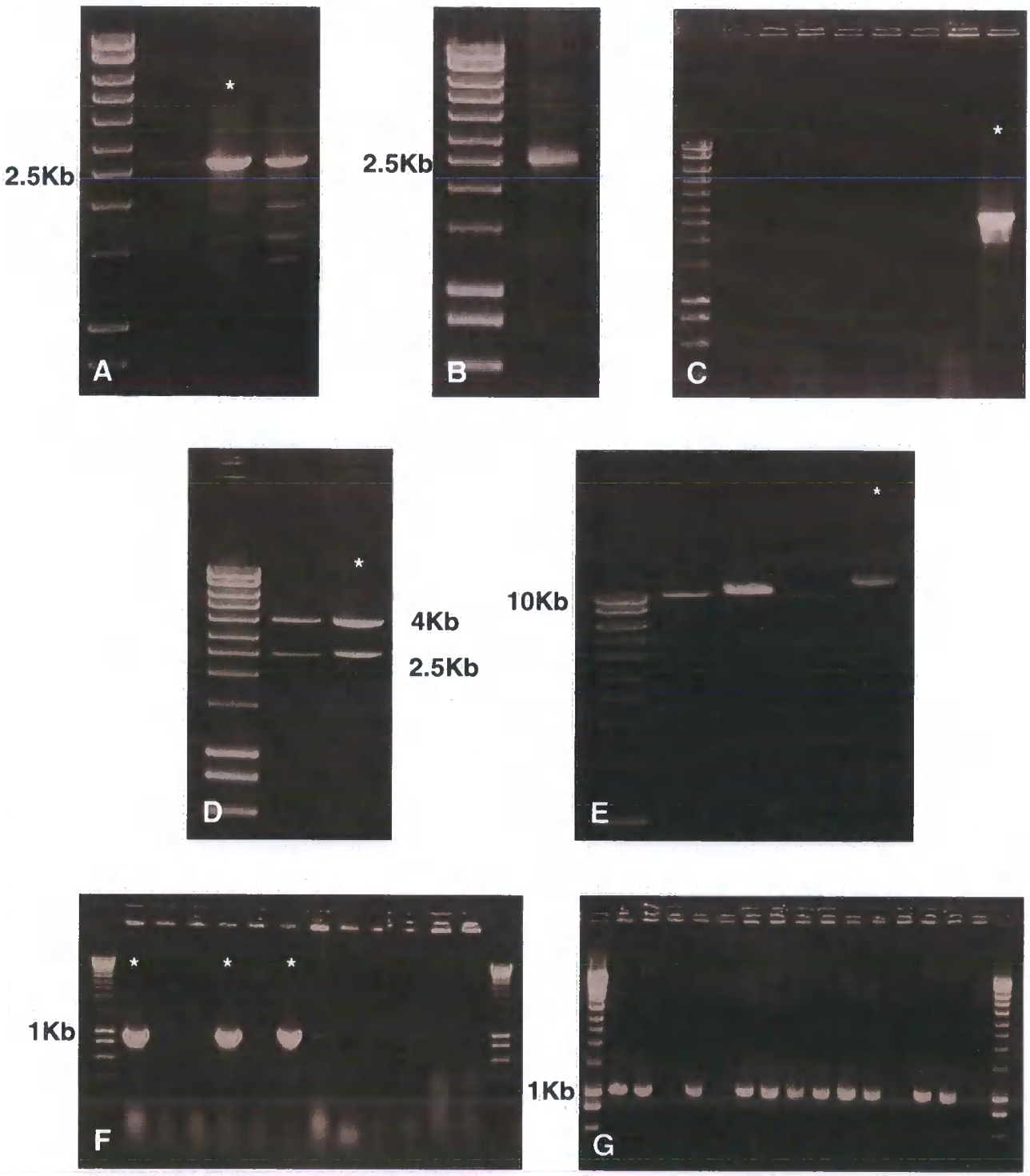

2.5Kb
A
2.5Kb
B
C
4Kb
2.5Kb
D
10Kb
E
1Kb
F
1Kb
G

Figure 3 Cloning of GUS Construct C (At5g14610)

A. Promoter PCR – Outer Reaction

Asterisked lane: ~2.5 kb promoter PCR fragment (no restriction sites).

B. Promoter PCR – Inner Reaction (addition of restriction sites)

Asterisked lane: ~2.5 kb promoter PCR fragment (BamH I restriction sites).

C. Colony PCR for promoter-TOPO colonies

Asterisked lanes: Putative positive colonies (~2.5 kb product).

D. Promoter fragment digested with BamH I (prior to purification)

Asterisked lane: ~2.5 kb promoter fragment and ~4 kb linearised TOPO vector.

E. pΔGUS-CIRCE digested with BamH I

Lanes 2-4: BamH I digested pΔGUS-CIRCE vector.

Lane 5 (asterisked): Un-digested pΔGUS-CIRCE vector.

F. Colony PCR for Promoter-pΔGUS-CIRCE colonies

Asterisked lanes: Putative positive colonies (~800 bp product).

G. Colony PCR for successfully transformed agrobacterium colonies

Asterisked lanes: Putative positive colonies (~800 bp product).

In all cases, lane 1 is Hyperladder I (Bioline), 10 kb – 100 bp.

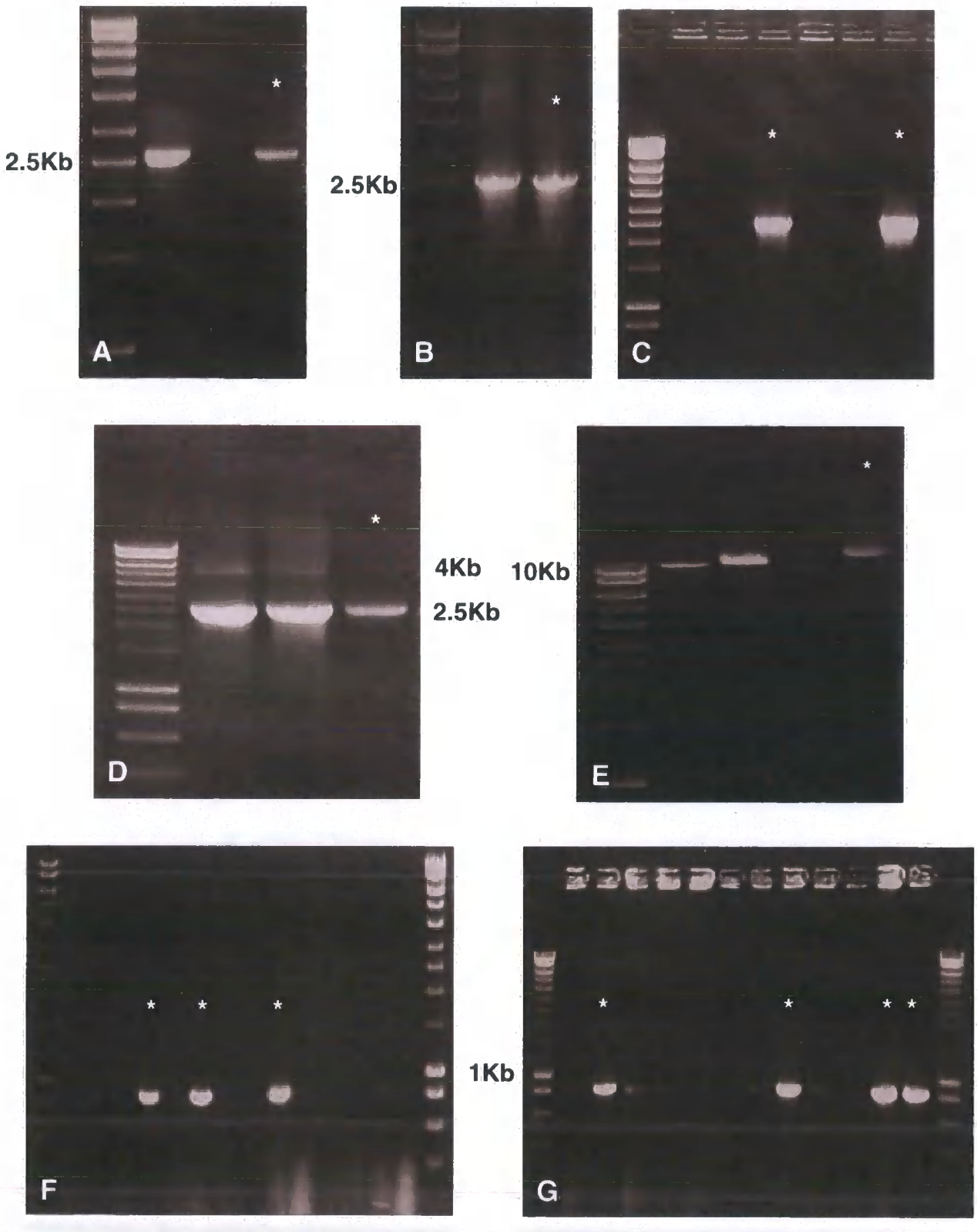

2.5Kb
A
2.5Kb
B
C
D
4Kb
2.5Kb
10Kb
E
F
1Kb
G

Figure 4 Cloning of GUS Construct D (At5g43040)

A. Promoter PCR – Outer Reaction

Asterisked lane: ~2.5 kb promoter fragment (no restriction sites).

B. Promoter PCR – Inner Reaction (addition of restriction sites)

Asterisked lane: ~2.5 kb promoter fragment (BamH I restriction sites).

C. Colony PCR for promoter-TOPO colonies

Asterisked lane: Putative positive colony (~2.5 kb product).

D. Promoter fragment digested with BamH I (prior to purification)

Asterisked lane: ~2.5 kb promoter fragment and ~4 kb linearised TOPO vector.

E. pΔGUS-CIRCE digested with BamH I

Lanes 2-4: BamH I digested pΔGUS-CIRCE vector.

Lane 5 (asterisked): Un-digested pΔGUS-CIRCE vector.

F. Colony PCR for Promoter-pΔGUS-CIRCE colonies

Asterisked lanes: Putative positive colonies (~800 bp product).

G. Colony PCR for successfully transformed agrobacterium colonies

Asterisked lanes: Putative positive colonies (~800 bp product).

In all cases the hyperladder used is Hyperladder I (Bioline), 10 kb – 100 bp.

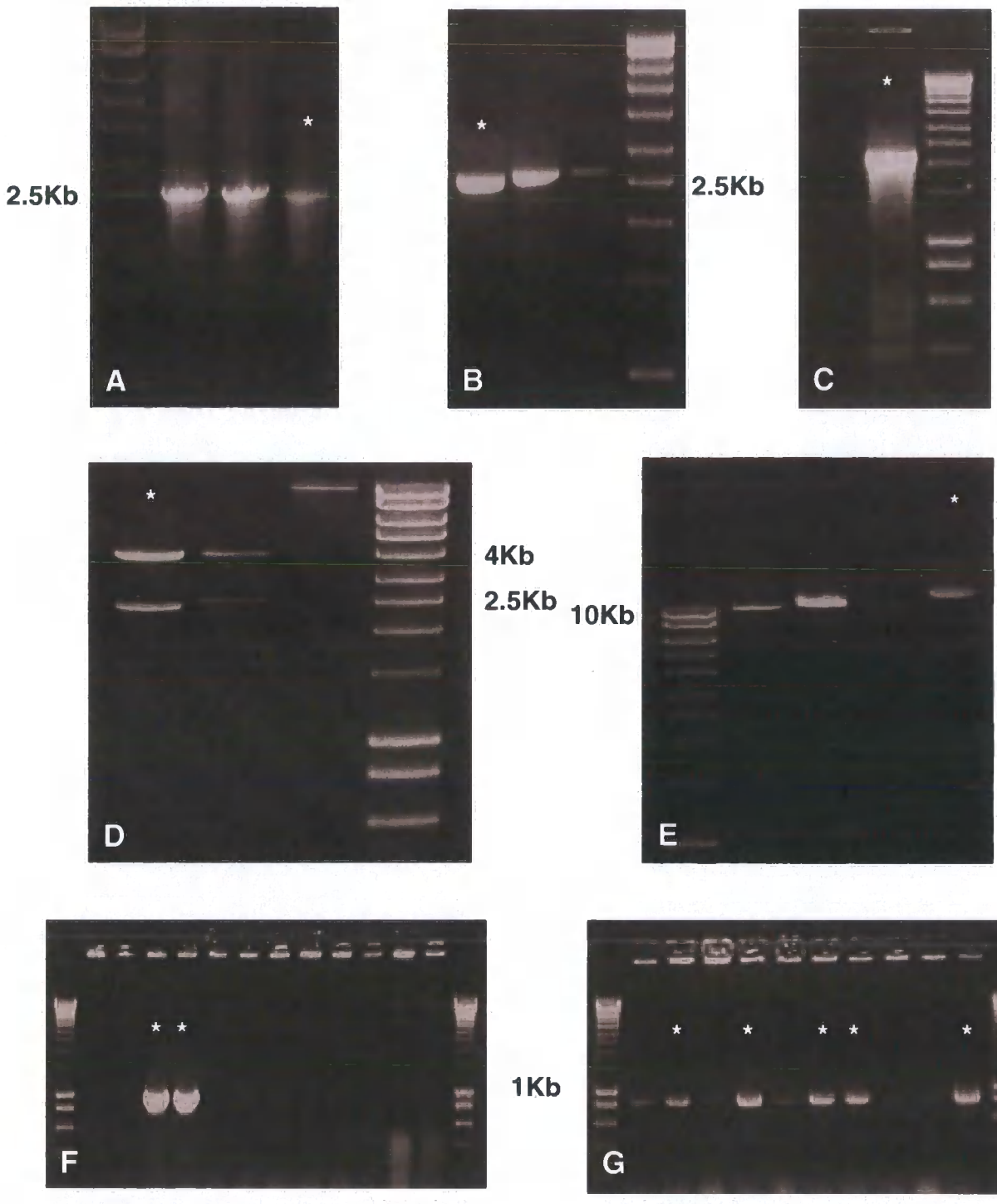

2.5Kb
A
B
2.5Kb
C
4Kb
2.5Kb
D
10Kb
E
1Kb
F
G

Figure 5 Cloning of GUS Construct E (At5g50810)

A. **Promoter PCR – Outer Reaction**
Asterisked lane: ~2.5 kb promoter fragment (no restriction sites).

B. **Promoter PCR – Inner Reaction (addition of restriction sites)**
Asterisked lane: ~2.5 kb promoter fragment (BamH I restriction sites).

C. **Colony PCR for promoter-TOPO colonies**
Asterisked lanes: Putative positive colonies (~2.5 kb product).

D. **Promoter fragment digested with BamH I (prior to purification)**
Asterisked lane: ~2.5 kb promoter fragment and ~4 kb linearised TOPO vector.

E. **pΔGUS-CIRCE digested with BamH I**
Lanes 2-4: BamH I digested pΔGUS-CIRCE vector.
Lane 5 (asterisked): Un-digested pΔGUS-CIRCE vector.

F. **Colony PCR for Promoter-pΔGUS-CIRCE colonies**
Asterisked lane: Putative positive colony (~800 bp product).

G. **Colony PCR for successfully transformed agrobacterium colonies**
Asterisked lanes: Putative positive colonies (~800 bp product).

In all cases the hyperladder used is Hyperladder I (Bioline), 10 kb – 100 bp.

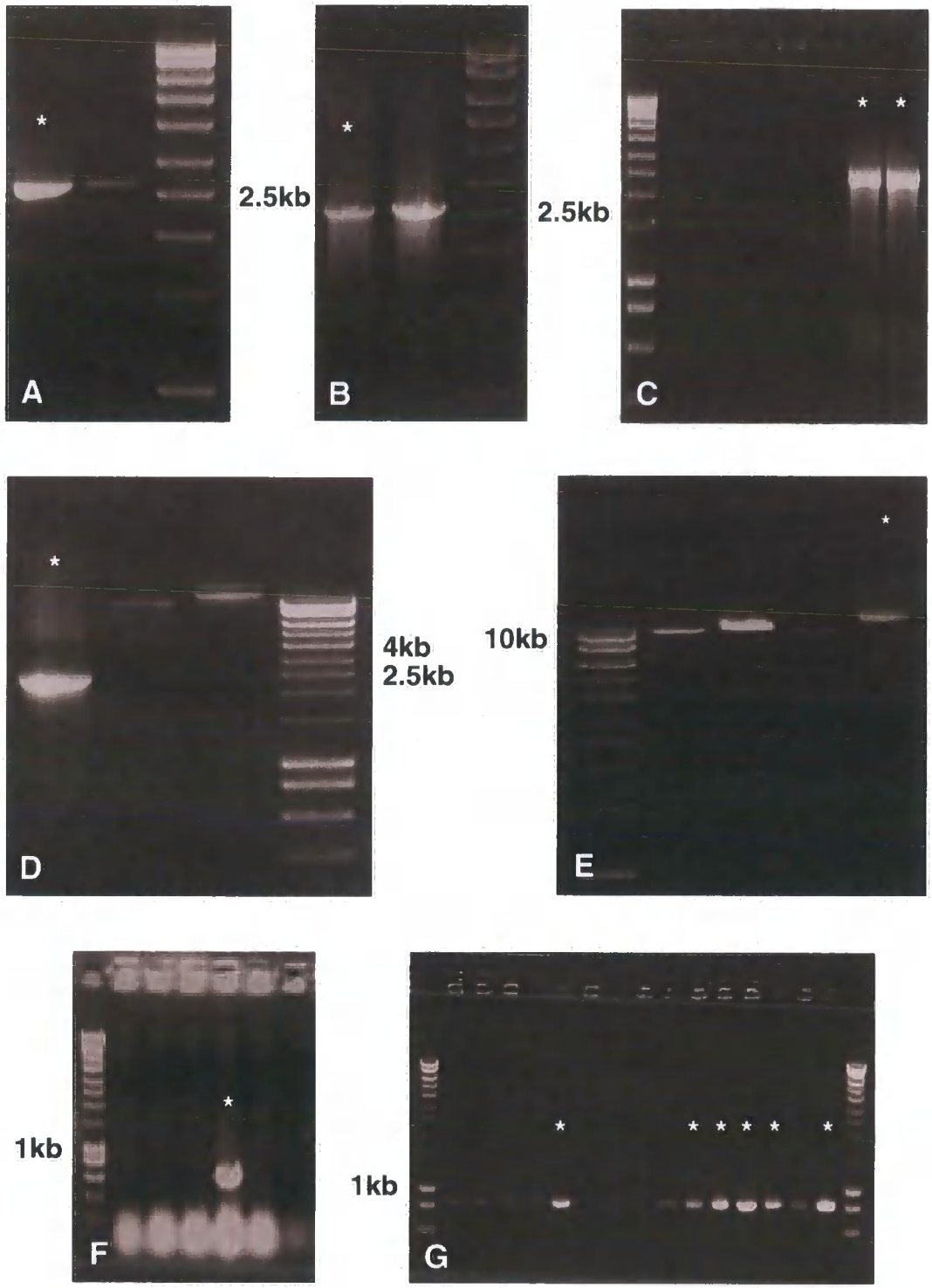

2.5kb
2.5kb
A
B
C
4kb
2.5kb
10kb
D
E
1kb
1kb
F
G

Figure 6 Cloning of GUS Construct F (At2g45050)

A. Promoter PCR – Outer Reaction

Lanes 1-3: ~2.5 kb promoter fragment (no restriction sites).

B. Promoter PCR – Inner Reaction (addition of restriction sites)

Lane 1: ~2.5 kb promoter fragment (BamH I restriction sites).

C. Colony PCR for promoter-TOPO colonies

Lanes 2, 4-8, 10-16: Putative positive colonies (~2.5 kb product).

D. Promoter fragment digested with BamH I (prior to purification)

Lane 1: ~2.5 kb promoter fragment, ~4 kb linearised TOPO vector and undigested vector.

In all cases the hyperladder used is Hyperladder I (Bioline), 10 kb – 100 bp.

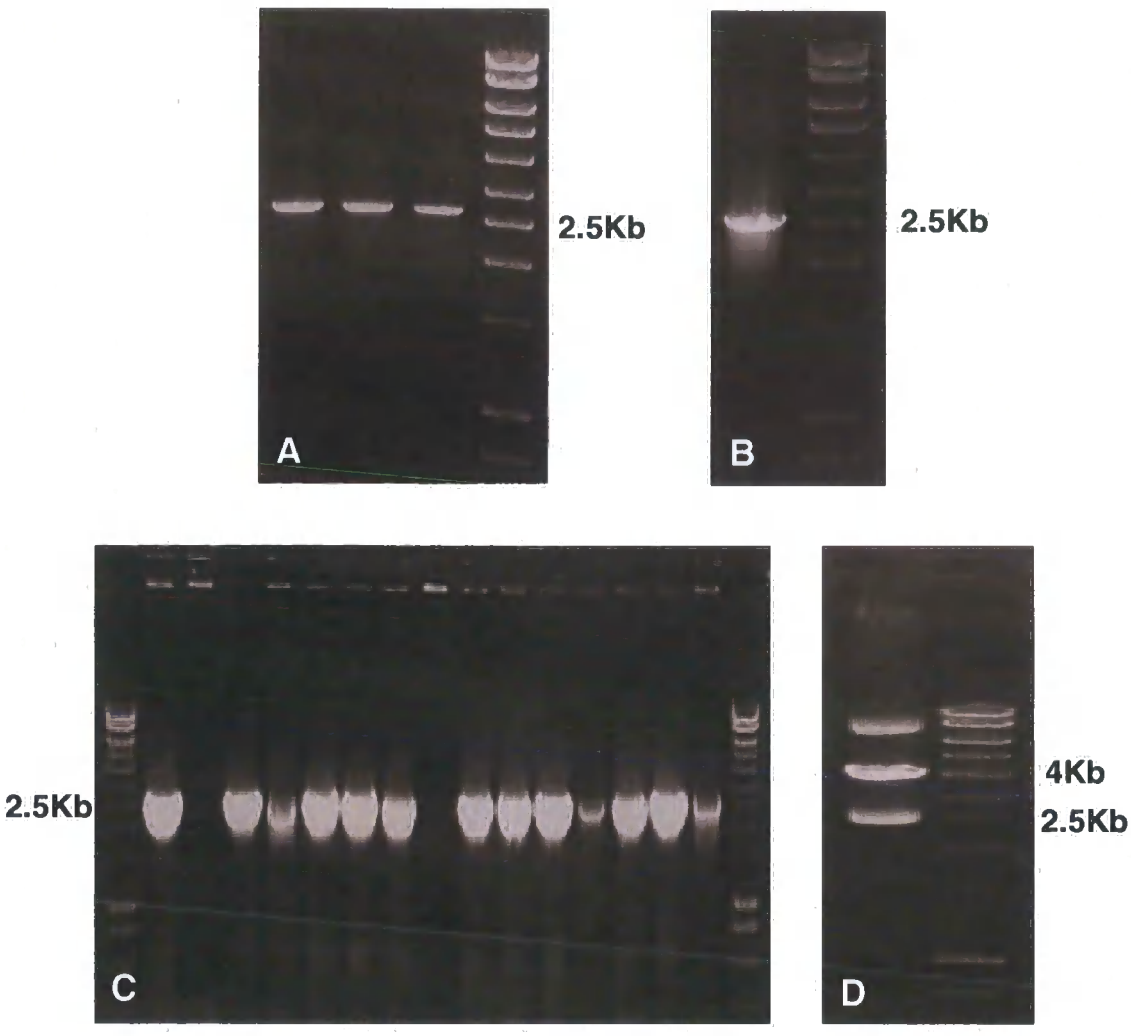

2.5Kb
A
2.5Kb
B
2.5Kb
C
4Kb
2.5Kb
D

Figure 7 Cloning of GUS Construct Q (At2g25420)

A. Promoter PCR – addition of Sal I restriction sites

Asterisked lane: ~2.5 kb promoter fragment (Sal I restriction sites).

B. Colony PCR for promoter-TOPO colonies

Asterisked lanes: Putative positive colonies (~2.5 kb product).

C. Promoter fragment digested with Sal I (prior to purification)

Asterisked lane: ~2.5 kb promoter fragment and ~4 kb linearised TOPO vector.

D. pΔGUS-CIRCE digested with Sal I

Lane 1 (asterisked): Un-digested pΔGUS-CIRCE vector.

Lanes 2-4: Sal I digested pΔGUS-CIRCE vector.

In all cases the hyperladder used is Hyperladder I (Bioline), 10 kb – 100 bp.

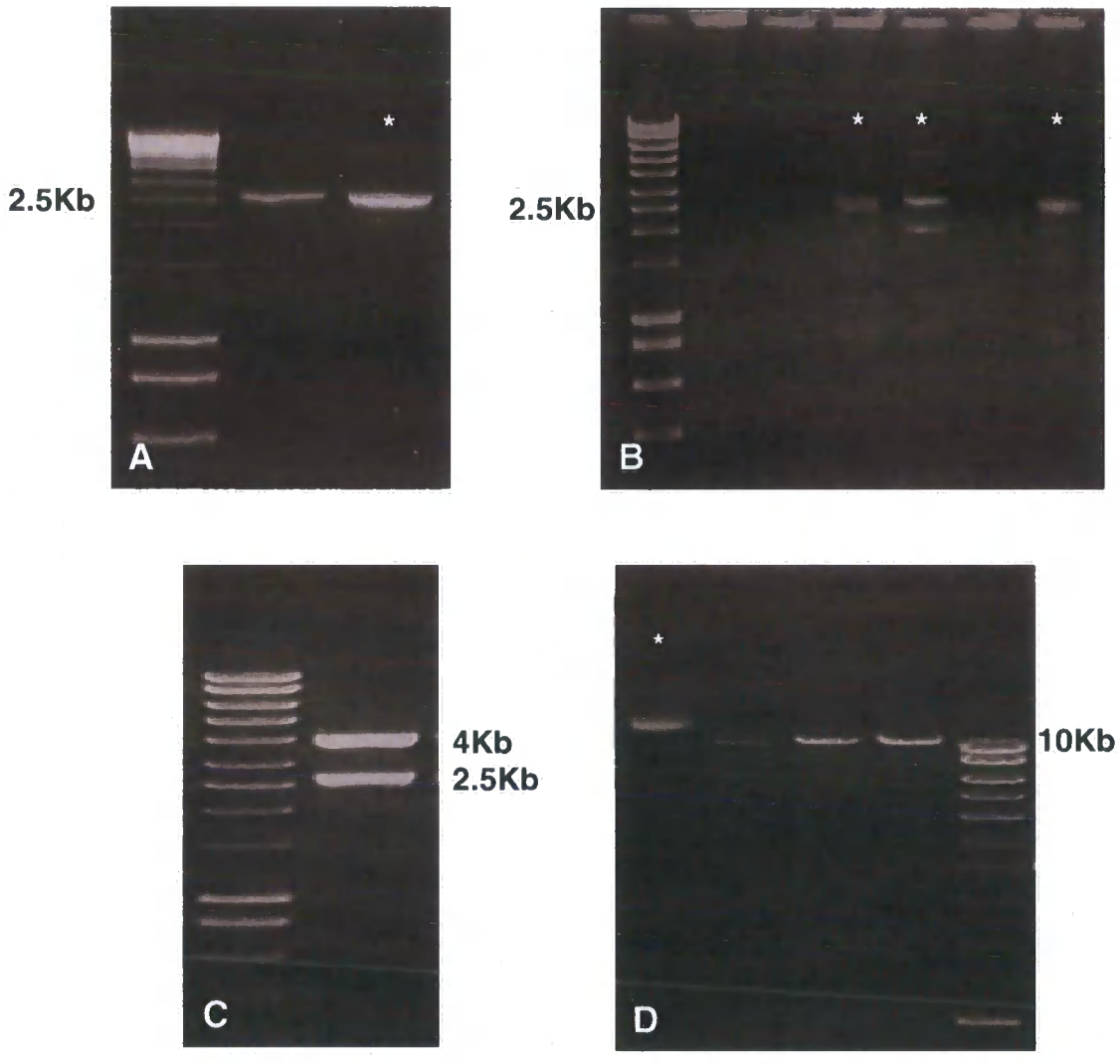

2.5Kb
A
2.5Kb
B
4Kb
2.5Kb
C
10Kb
D

Figure 8 Cloning of GUS Construct R (At2g31510)

A. Promoter PCR – addition of BamH I restriction sites

Asterisked lanes: ~2.5 kb promoter fragments (Bam HI restriction sites).

B. Colony PCR for promoter-TOPO colonies

Asterisked lane: Putative positive colony (~2.5 kb product).

C. Promoter fragment digested with BamH I (prior to purification)

Asterisked lane: ~2.5 kb promoter fragment and ~4 kb linearised TOPO vector.

D. pΔGUS-CIRCE digested with BamH I

Lanes 2-4: BamH I digested pΔGUS-CIRCE vector.

Lane 5 (asterisked): Un-digested pΔGUS-CIRCE vector.

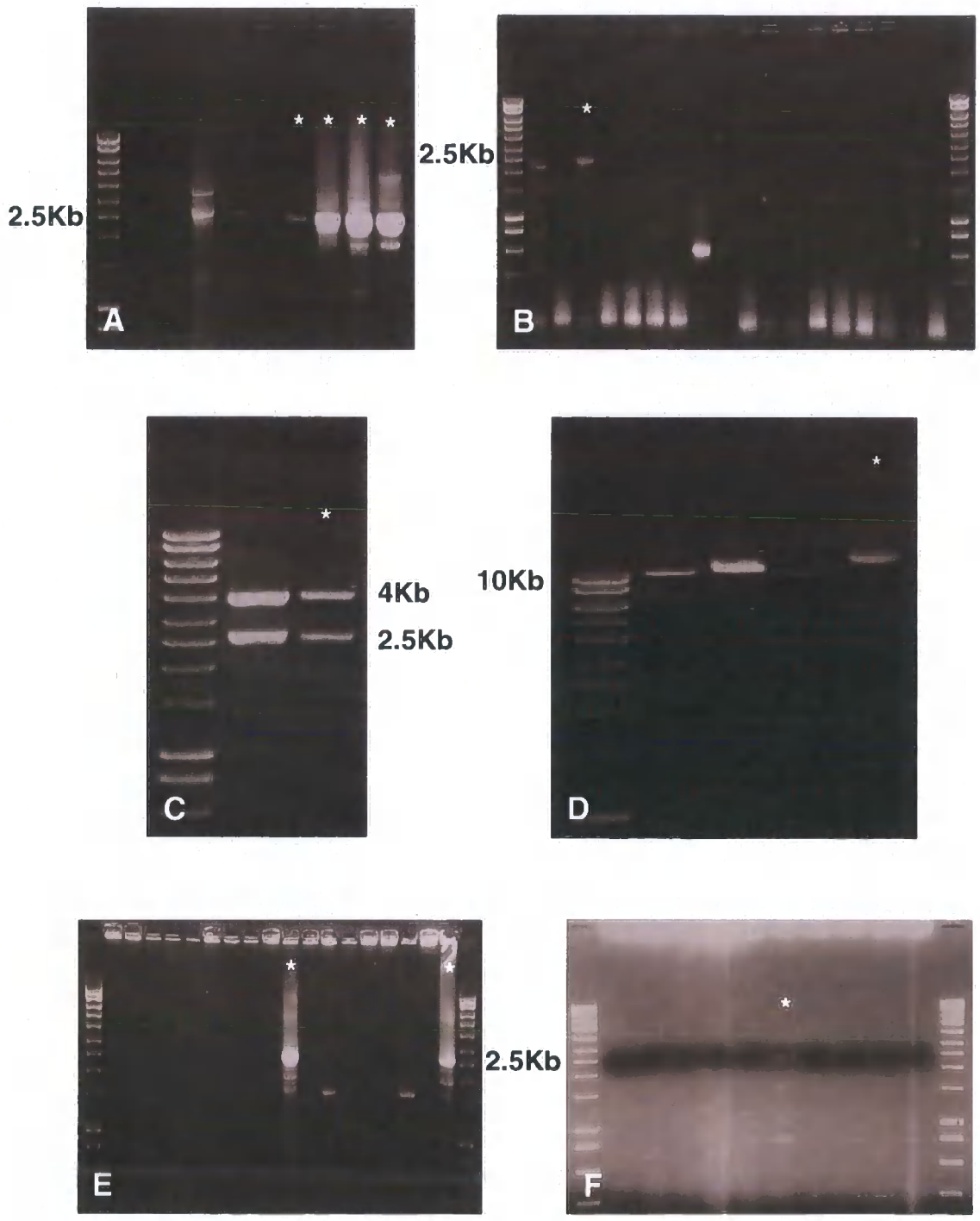
E. Colony PCR for Promoter-pΔGUS-CIRCE colonies

Asterisked lanes: Putative positive colonies (~2.5 kb product).

F. Colony PCR for successfully transformed agrobacterium colonies

Asterisked lane: Putative positive colony (~2.5kb product).

In all cases the hyperladder used is Hyperladder I (Bioline), 10 kb – 100 bp.


2.5Kb
2.5Kb
A
B
4Kb
2.5Kb
C
10Kb
D
2.5Kb
E
F

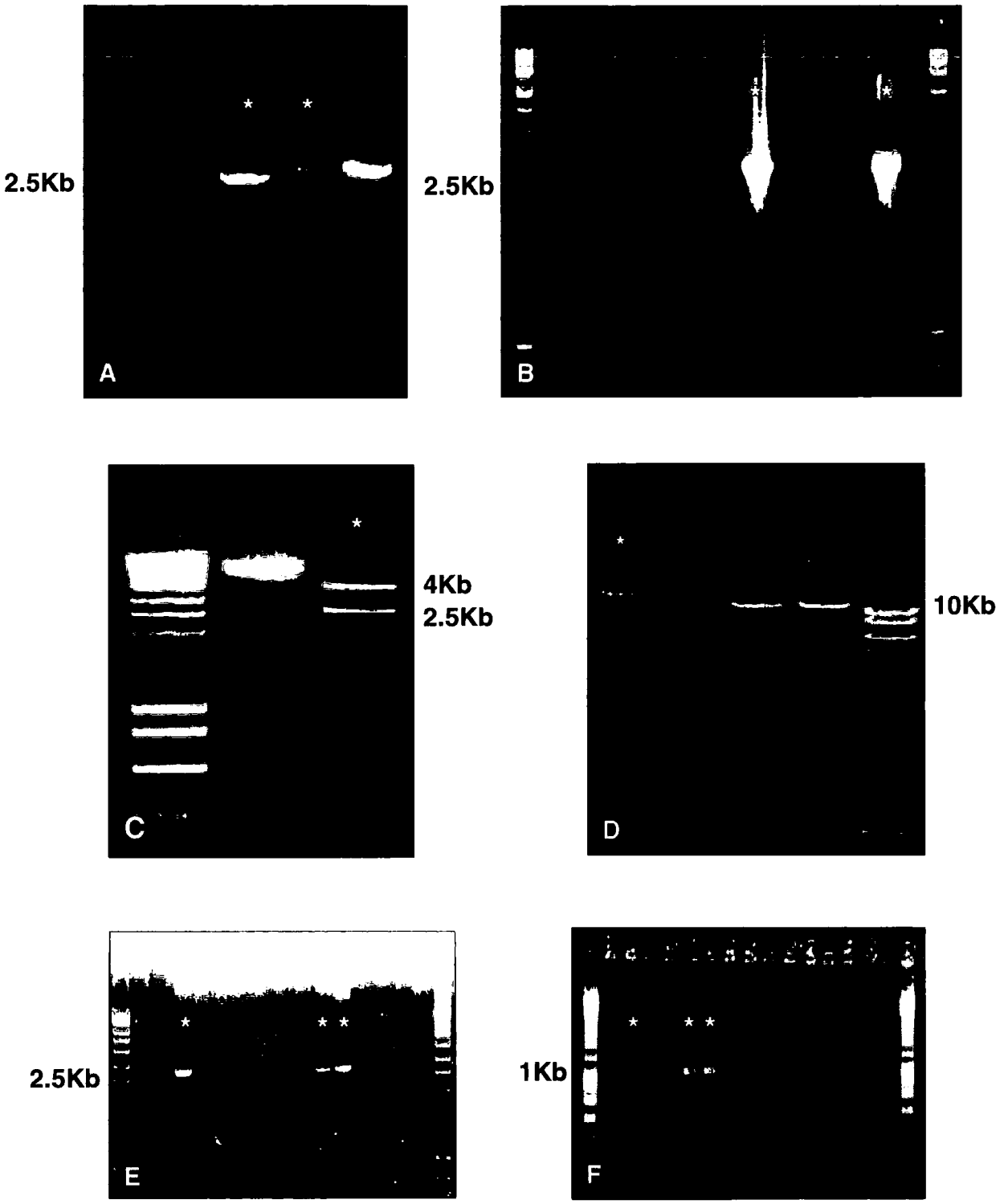

2.5Kb
A
2.5Kb
B
4Kb
2.5Kb
C
10Kb
D
2.5Kb
E
1Kb
F

Figure 10 Cloning of GUS Construct T (At5g45600)

A. Promoter PCR

Asterisked lane: ~2.5 kb promoter fragment.

B. Colony PCR for promoter-TOPO colonies

Asterisked lanes: Putative positive colonies (~2.5 kb product).

C. Promoter fragment digested with BamH I and Xba I

Asterisked lane: ~2.5 kb promoter fragment and ~4 kb linearised TOPO vector. Subsequently 3' adenine residues were added to promoter fragment.

D. pΔGUS-CIRCE digested with Sma I

Lane 3 (asterisked): Un-digested pΔGUS-CIRCE vector.

Lane 4: Sma I digested pΔGUS-CIRCE vector. Digested vector T-tailed to allow ligation.

E. Colony PCR for Promoter-pΔGUS-CIRCE colonies

Asterisked lane: Putative positive colony (~2.5 kb product).

F. Colony PCR for successfully transformed agrobacterium colonies

Lanes 2-5, 7, 10-11: Putative positive colonies (~2.5 kb product).

In all cases the hyperladder used is Hyperladder I (Bioline), 10 kb – 100 bp.

Figure 9 Cloning of GUS Construct S (At3g60860)

A. Promoter PCR – addition of Sal I restriction sites

Asterisked lanes: ~2.5 kb promoter fragments (Sal I restriction sites).

B. Colony PCR for promoter-TOPO colonies

Asterisked lanes: Putative positive colonies (~2.5 kb product).

C. Promoter fragment digested with Sal I (prior to purification)

Asterisked lane: ~2.5 kb promoter fragment and ~4 kb linearised TOPO vector.

D. pΔGUS-CIRCE digested with Sal I

Lane 1 (asterisked): Un-digested pΔGUS-CIRCE vector.

Lanes 2-4: Sal I digested pΔGUS-CIRCE vector.

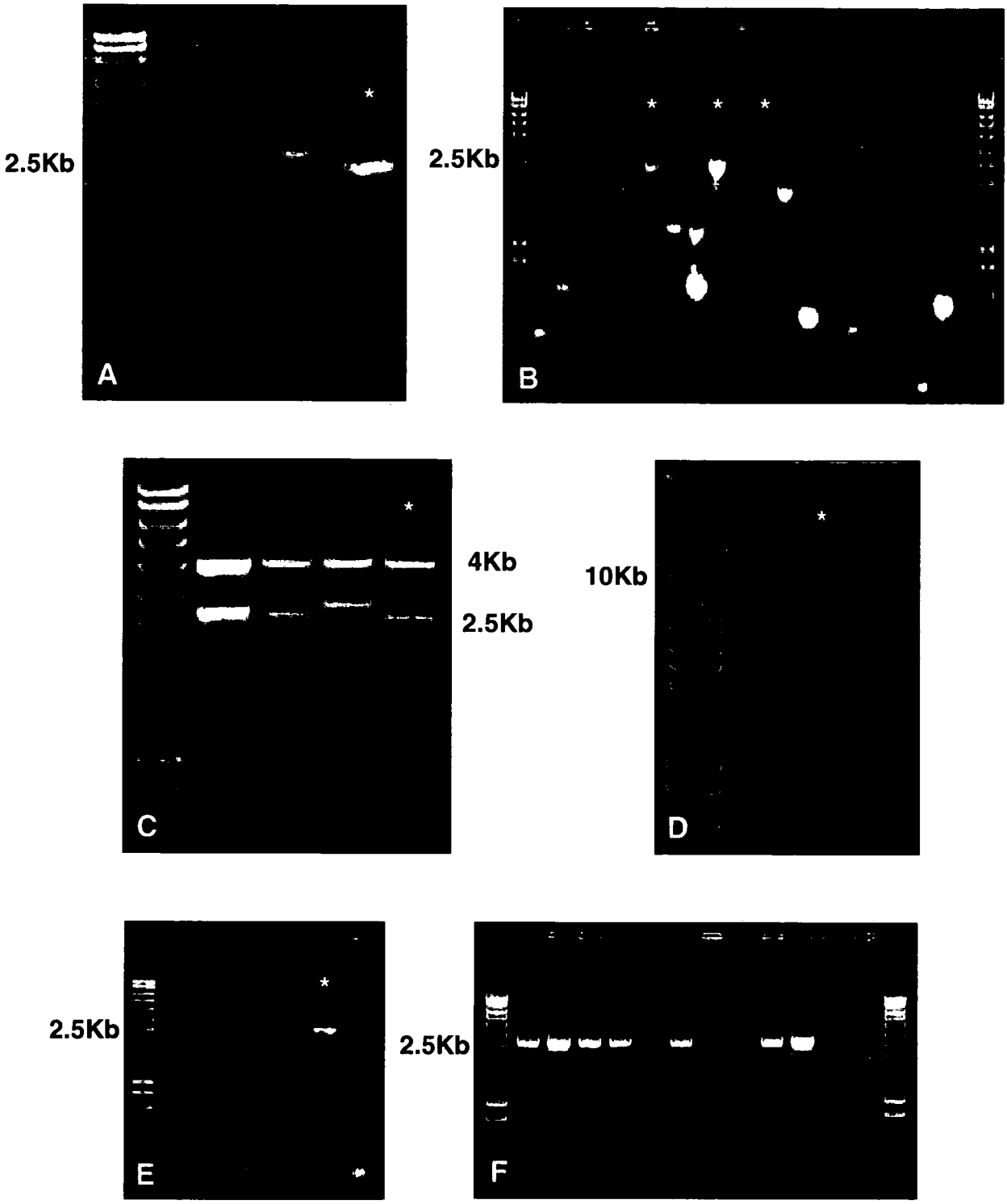
E. Colony PCR for Promoter-pΔGUS-CIRCE colonies

Asterisked lanes: Putative positive colonies (~2.5 kb product).

F. Colony PCR for successfully transformed agrobacterium colonies

Asterisked lanes: Putative positive colonies (~2.5 kb product).

In all cases the hyperladder used is Hyperladder I (Bioline), 10 kb – 100 bp.


2.5Kb
A
2.5Kb
B
4Kb
2.5Kb
C
10Kb
D
2.5Kb
E
2.5Kb
F